THE DESIGN AND ANALYSIS OF PEDIATRIC VACCINE FORMULARIES: THEORY AND PRACTICE

BY

SHANE NIKOLAUS HALL

B.S., Brigham Young University, 1997
M.S., Air Force Institute of Technology, 2000

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Industrial Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2006

Urbana, Illinois

Report Documentation Page

Form Approved
OMB No. 0704-0188

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE **22 JUN 2006**	2. REPORT TYPE **N/A**	3. DATES COVERED **-**
4. TITLE AND SUBTITLE **The Design And Analysis Of Pediatric Vaccine Formularies: Theory And Practice**		5a. CONTRACT NUMBER 5b. GRANT NUMBER 5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)		5d. PROJECT NUMBER 5e. TASK NUMBER 5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **University of Illinois at Urbana-Champaign**		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)		10. SPONSOR/MONITOR'S ACRONYM(S) 11. SPONSOR/MONITOR'S REPORT NUMBER(S)
12. DISTRIBUTION/AVAILABILITY STATEMENT **Approved for public release, distribution unlimited**		
13. SUPPLEMENTARY NOTES		
14. ABSTRACT		
15. SUBJECT TERMS		

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**UU**	**152**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

CERTIFICATE OF COMMITTEE APPROVAL

University of Illinois at Urbana-Champaign
Graduate College

June 22, 2006

We hereby recommend that the thesis by:

SHANE NIKOLAUS HALL

Entitled:

THE DESIGN AND ANALYSIS OF PEDIATRIC VACCINE FORMULARIES: THEORY AND PRACTICE

Be accepted in partial fulfillment of the requirements for the degree of:

Doctor of Philosophy

Signatures:

Director of Research - Sheldon H. Jacobson

Head of Department - Huseyin Sehitoglu

Committee on Final Examination*

Chairperson - Sheldon H. Jacobson

Committee Member - Vinayak Uday Shanbhag

Committee Member - Udatta S. Palekar

Committee Member -

Committee Member - Edward C. Sewell

Committee Member -

*R equired for doctoral degree but not for master's degree

Abstract

Vaccination against infectious disease is hailed as one of the greatest public health achievements. However, the United States Recommended Childhood Immunization Schedule is becoming increasingly complex, often requiring numerous separate injections in a single pediatric visit. To address the issue of vaccine delivery complexity, vaccine manufacturers have developed combination vaccines that immunize against several diseases in a single injection. These combination vaccines are creating challenges such as how these vaccines should be administered to ensure that immunity is safely achieved. Furthermore, these vaccines are also creating a combinatorial explosion of alternatives and choices for public health policy-makers and administrators, pediatricians, and parents/guardians.

This dissertation applies operations research methodologies to designing pediatric vaccine formularies that capture this combinatorial explosion of alternatives and choices and ensure that immunity is safely achieved. In particular, the dissertation presents three fundamental problems for designing pediatric vaccine formularies.

The first problem models a general childhood immunization schedule to design a vaccine formulary that minimizes the cost of fully immunizing a child. The second problem models a general childhood immunization schedule to design a vaccine formulary that safely immunizes a child against several infectious diseases by restricting or limiting extraimmunization (i.e., extra doses of vaccine). These problems are vitally important since the cost of vaccinating a child contributes to the underimmunization of children, and extraimmunization poses biological risks, amplifies philosophical concerns with vaccination, and creates an unnecessary economic burden on society. These models are rigorously analyzed and several algorithms—both exact and heuristic—are presented. Furthermore, a computational comparison of these algorithms is presented for the 2006 Recommended Childhood Immunization Schedule as well as several randomly generated childhood immunization schedules. The third problem combines the first two problems by modeling a general childhood immunization schedule to design a vaccine formulary that minimizes the cost of fully immunizing a child while restricting or limiting extraimmunization. The results reported here provide both fundamental insights to the operations research community as well as practical value for the public health community.

For my two children who have enriched my life with love and joy beyond expression or quantification.

Acknowledgements

This dissertation represents in some measure my attempt to contribute in a small way to the field of operations research. This effort has required countless hours of reading, thinking, writing, coding, questioning, and analyzing, and has been filled with moments of excitement, appreciation, and wonderment, as well as worry, anxiety, and exhaustion. This acknowledges those individuals and institutions (in no particular order) that have unselfishly sacrificed their time and resources, and have supported me throughout my life and during my graduate studies at the University of Illinois.

I sincerely appreciate the influence and guidance of my academic mentor and advisor, Professor Sheldon H. Jacobson. I admire his intelligence, enthusiasm, and drive for excellence. He was always positive, ever encouraging, and exceedingly patient. I look forward to future years of collaboration and friendship. I also appreciate the additional guidance and support from my research committee: Professors Edward C. Sewell, Udatta S. Palekar, Uday V. Shanbhag, Diego Klabjan, and Joseph P. Goldberg. I respect and admire each of these professors and feel honored by their association.

I also express my appreciation for the many exceptional teachers, both here at the University of Illinois and in years past, which have inspired me to use logic, think critically, appreciate knowledge, and pursue truth. I also thank the entire University of Illinois community and the Department of Mechanical and Industrial Engineering faculty, staff, and fellow students for an outstanding educational experience. I am also appreciative to the Department of Mechanical and Industrial Engineering for naming me as the 2006 George W. Harper Award recipient. I also thank the 2004-05 Fighting Illini men's basketball team for a truly miraculous season and converting me to an Illini orange-wearing fan!

I am honored to be a commissioned officer in the United States Air Force, and express my appreciation to the Air Force for allowing me the time and financial support to pursue my education. I am thankful for their commitment to education and professional excellence. I express my thanks for the trust and confidence that the faculty and staff of the Department of Operational Sciences, Air Force Institute of Technology have placed in me by sponsoring this educational experience. I wish to thank Professors Jim Moore and Ray Hill for their guidance on my Masters thesis and for their encouragement to pursue a Ph.D degree. It has been a privilege to share this assignment with other Air Force officers and enlisted troops stationed here at the University of Illinois. I thank them for their support and friendship.

I extend my thanks to the National Science Foundation (DMI-0222597 and DMI-0457176) and the Air Force Office of Scientific Research (FA9550-04-1-0110) that provided funding for this research. The views expressed in this dissertation are mine and do not reflect the official policy or position of the United States Air Force, Department of Defense, National Science Foundation, or the United States Government.

On a more personal note, I must also acknowledge the many exceptional individuals that I have had the privilege to meet in Illinois. To my research group, Gio Kao, Hemanshu Kaul, Adrian Lee, Laura McKlay, Alex Nikolaev, and Ruben Proano, I express a sincere thanks for their friendship and insightful conversations on topics such as mathematics, complexity theory, philosophy, religion, and politics. To my running group, Pat Castle, John Shannon, and Bruce Flaschbart, I say thanks for helping me stay in shape, both mentally and physically, and finally, to the many other devoted friends like Ryan Hanna and Kirk Price that have enriched my life more than they realize.

Finally, I must especially acknowledge the unfailing love and support of my parents and to my grandparents whose lessons, lives, and love have profoundly influenced me. To my siblings, I also thank for their influence and love. To my two children, I express my love and gratitude for the inspiration and joy they render in me and thank them for their sacrifice during these years when my time and attention were largely consumed by school.

The views expressed in this article are those of the author and do not reflect the official policy or position of the United States Air Force, Department of Defense, or the U.S. Government.

Table of Contents

List of Tables

Page

List of Figures

Chapter 1: Motivation and Introduction

Infectious diseases have plagued humankind for centuries. Even as recently as the early 20th century, infectious disease was the leading worldwide killer (Cohen 2000). For example, in 1900, the average life expectancy at birth in the United States was 47 years, with nearly thirty percent of all deaths caused by infectious disease. Moreover, roughly one in ten children died before the age of four due to infectious diseases (Cohen 2000).

By the end of the 20th century, the life expectancy in the United States increased to over 76 years (National Vital Statistics Reports 2004). Naturally, it is worth considering how life expectancy in this country has nearly doubled over the past century. According to the World Health Organization (WHO), vaccinations and clean water have had the greatest impact on world health (Plotkin and Orenstein 2004). As evidence to this claim, immunization spares millions of children each year from contracting infectious diseases (Cohen 2000, Diekema 2005). Moreover, the United States has witnessed a 100% decrease in the number of cases of indigenous poliomyelitis, and a 99% decrease in the number of cases of diphtheria, measles, mumps, and rubella since vaccines became available. Furthermore, in 1966, there were an estimated 20M cases of smallpox worldwide, and because of vaccination, the WHO declared the eradication of this disease in 1980 (Mackay and Rosen 2001). Today many healthcare professionals still regard the eradication of smallpox as one of the greatest accomplishments of public health (Cohen 2000, Mackay and Rosen 2001).

In spite of this progress against infectious disease, much work still remains. For example, in 1998, almost a quarter of the deaths worldwide (over 13M of the 54M deaths) were caused by infectious disease (Cohen 2000). Unfortunately, an estimated 1M of these deaths were attributed to measles alone—a disease with a readily available vaccine at a relatively inexpensive cost of US$15 per dose (Cohen 2000, Plotkin and Orenstein 2004). The emergence of new infectious diseases such as human immunodeficiency virus (HIV) and Lyme disease, the resurgence of diseases such as tuberculosis, and the recent threat of bioterrorism (anthrax, smallpox, etc.) highlight the need for continued vigilance in the effort to combat infectious diseases (Binder et al. 1999, Plotkin and Orenstein 2004).

1.1 Pediatric Immunization in the United States

The National Immunization Program (NIP) within the United States Centers for Disease Control and Prevention (CDC) is the federal agency responsible for overseeing pediatric immunization practice and policy. Each year, based on recommendations from the Advisory Committee on Immunization Practices (ACIP) and the American Academy of Family Physicians (AAFP), the NIP publishes a Recommended Childhood Immunization Schedule that outlines vaccination requirements for children through adolescence (CDC 2006). The Recommended Childhood Immunization Schedule (see Figure 1) outlines the vaccinations required to protect a child against several (currently thirteen) infectious diseases that pose a risk to children living in the United States. This schedule includes the number of required doses of each vaccine and the recommended age for each dose (D1 = Dose 1, D2 = Dose 2, etc.). For example, polio requires four doses of vaccine, where the third dose (D3) may be administered at age 6 months, 12 months, 15 months, or 18 months.

<table>
<tr><th rowspan="2">DISEASE</th><th colspan="10">TIME PERIOD (Age of Child)</th></tr>
<tr><th>1
(Birth)</th><th>2
(1 Mo)</th><th>3
(2 Mos)</th><th>4
(4 Mos)</th><th>5
(6 Mos)</th><th>6
(12 Mos)</th><th>7
(15 Mos)</th><th>8
(18 Mos)</th><th>9
(24 Mos)</th><th>10
(4-6 Yrs)</th></tr>
<tr><td>Hepatitis B</td><td>D1</td><td colspan="2">D2</td><td></td><td colspan="4">D3</td><td></td><td></td></tr>
<tr><td>Diphtheria, Tetanus, Pertussis</td><td></td><td></td><td>D1</td><td>D2</td><td>D3</td><td></td><td colspan="2">D4</td><td></td><td>D5</td></tr>
<tr><td>Haemophilus influenzae type b</td><td></td><td></td><td>D1</td><td>D2</td><td>D3</td><td colspan="2">D4</td><td></td><td></td><td></td></tr>
<tr><td>Polio</td><td></td><td></td><td>D1</td><td>D2</td><td colspan="4">D3</td><td></td><td>D4</td></tr>
<tr><td>Measles, Mumps, Rubella</td><td></td><td></td><td></td><td></td><td></td><td colspan="2">D1</td><td></td><td></td><td>D2</td></tr>
<tr><td>Varicella</td><td></td><td></td><td></td><td></td><td></td><td colspan="3">D1</td><td></td><td></td></tr>
<tr><td>Pneumococcus</td><td></td><td></td><td>D1</td><td>D2</td><td>D3</td><td colspan="2">D4</td><td></td><td></td><td></td></tr>
<tr><td>Influenza</td><td></td><td></td><td></td><td></td><td colspan="6">D1 (yearly)</td></tr>
<tr><td>Hepatitis A</td><td></td><td></td><td></td><td></td><td></td><td>D1</td><td></td><td>D2</td><td></td><td></td></tr>
</table>

Figure 1: United States 2006 Recommended Childhood Immunization Schedule through Age 6 (excluding recommendations for selected populations)

Each vaccine dose is typically administered by injection during a scheduled well-baby check-up at a health care clinic. For example, an infant child should receive a dose of vaccine for hepatitis B, diphtheria, tetanus, pertussis, *Haemophilus influenzae* type b, polio, and pneumococcus at their two-month well-baby check-up, resulting in as many as five injections. Furthermore, a fifteen-month old child, under extreme conditions, could receive as many as eight injections in a single clinic visit. These examples demonstrate that the

Recommended Childhood Immunization Schedule is becoming overly crowded and complex. Moreover, this situation will only worsen in the future as new diseases emerge and/or new vaccines are developed. For example, four time periods and three diseases have been added to the Recommended Childhood Immunization Schedule since 1995, and there are currently several vaccine products being marketed and tested for use in children (CDC 1995, Cochi 2005, Infectious Diseases in Children 2002). These added complexities in the Recommended Childhood Immunization Schedule increase the likelihood that a parent/guardian will reject or delay some vaccinations. The cost of vaccinating a child also contributes to the underimmunization of children—the opportunity cost of time for a parent/guardian to make clinic visits as well as the monetary cost of vaccination (Plotkin and Orenstein 2004). These costs often contribute to either missed clinic visits or missed vaccine doses. For example, the three year measles epidemic in the United States that began in 1990 involved 28,000 cases of measles, most of which were due to inadequate vaccination of these patients when they were one to two years of age (Mackay and Rosen 2001). One estimate is that pediatric immunizations prevent three million worldwide deaths in children each year (Diekema 2005), and hence, noncompliance to the Recommended Childhood Immunization Schedule puts children at risk of contracting potentially debilitating (and sometimes fatal) infectious diseases, thereby creating an enormous cost burden (both tangible and intangible) on the individual child, family, and society at large. For example, the CDC's 2005 National Immunization Survey estimates a savings of US$27 in direct and indirect costs for every dollar spent on vaccinating against diphtheria, tetanus, and pertussis (Cochi 2005).

Weniger (1996) discusses several options that address the issues of vaccine injection overcrowding, schedule complexity, and the cost of vaccinating a child. The most feasible option is the development and use of combination vaccines—a vaccine that combines several antigens (a substance that stimulates the production of an antibody, i.e., toxins, bacteria, foreign blood cells, and the cells of transplanted organs) into a single injection. Some combination vaccines are already commonly used, such as the DTaP vaccine, which combines diphtheria and tetanus toxoids with acellular pertussis vaccine. The ideal combination vaccine would combine antigens for every disease in the Recommended Childhood Immunization Schedule into a single vaccine, which could be administered at birth. However, developing such a vaccine is highly unlikely based on financial and biological constraints. For example, live vaccines (vaccines that inject living antigens) can interfere with each other by competing for binding sites. Nonetheless, several pediatric combination vaccines are now coming to market, and several more are being developed and tested for licensure in the United States (Infectious Diseases in Children 2002). For example, Proquad®, a combination vaccine manufactured by Merck that immunizes against measles, mumps, rubella, and varicella, gained Food and Drug Administration (FDA) approval in September 2005.

Combination vaccines will alleviate the issue of vaccine injection overcrowding and also offer economic opportunities by being more affordable per dose and reducing the shipping, handling, and storage costs of vaccines (Edwards and Decker 2001). However, combination vaccines also pose several challenging questions, such as which antigens should be combined and how should these vaccines be administered to ensure that immunity is safely achieved and remains economically reasonable. Moreover, combination vaccines offer pediatricians, public health policy-makers and administrators, and parents/guardians additional alternatives and choices on how to best immunize a child, and hence, these choices amplify the schedule complexity. In fact, as the Recommended Childhood Immunization Schedule continually evolves, new combination vaccines will lead to a combinatorial explosion of alternatives and choices for such individuals, each with a different cost. Therefore, determining the set of vaccines that minimize the cost of immunizing a child becomes more challenging.

Furthermore, combination vaccines increase the risk of extraimmunization. Extraimmunization means that a child receives antigens for a given disease over the recommended quantity and timing sequence. Since combination vaccines reduce the number of required injections and may be more economical, pediatricians, public health policy-makers and administrators, and parents/guardians will likely choose combination vaccines over multiple single antigen vaccines. However, using combination vaccines may inject a child with antigens they have already received in the recommended quantity and timing sequence. For example, injecting a child with a DTaP-HBV-IPV (diphtheria, tetanus, pertussis, hepatitis B, and polio) combination vaccine at age 4 months would provide extraimmunization for hepatitis B, since (according to Figure 1) no dose of vaccine is required at that age. Such extraimmunization poses biological risks and amplifies philosophical concerns. Biologically, extraimmunization of some antigens increases the risk of adverse side effects. Such is the case with diphtheria and tetanus vaccines (CDC 1999). Philosophically, many people challenge the safety and effectiveness of vaccinating children and particularly object to the use of combination vaccines, because they believe injecting a child with multiple antigens simultaneously overwhelms the infant immune system, and hence, extraimmunization due to combination vaccines only increases these fears (Edwards and Decker 2001, Chen et al. 2001). This philosophical barrier to vaccination is an increasing concern for pediatricians and public health administrators. For example, in a recent national survey of pediatricians, 54 percent had encountered parents over a 12-month period that refused to vaccinate their child, citing safety concerns as the top reason for this refusal (Flanagan-Klygis, Sharp and Frader 2005). In another survey, 70 percent of pediatricians had encountered a parent in the 12 months preceding the survey that refused at least one immunization for their child (Diekema 2005). In addition to these biological and philosophical concerns, the economic toll of extraimmunization is significant. For example, the annual societal cost burden of providing one extra dose of vaccine for each child born in the United States is over $28 million, which assumes a birth rate of 11,100 births per day (see Jacobson, et al. 2006a, 2006b) and a vaccine cost of $7, both of which are conservative estimates, where the vaccine cost estimates the Federal contract purchase price of the least expensive pediatric vaccine (see CDC Vaccine Price List 2005).

1.2 Dissertation Overview

This dissertation addresses how vaccines can best be administered to ensure immunity is safely achieved at a reasonable cost by examining the issues of cost and extraimmunization, and is organized as follows.

Chapter 2 presents a literature review of earlier research where operations research techniques have been used to address pediatric immunization problems. Other relevant topics related to this research are also discussed.

Chapter 3 presents general models (formulated as a decision problem and as a discrete optimization problem) that determine the set of vaccines (i.e., a vaccine formulary) that should be used in a clinical environment to satisfy any given childhood immunization at minimum cost, rigorously explores the theoretical structure of these general models, and provides an extensive computational study. Specifically, Chapter 3 presents the computational complexity of the decision/discrete optimization problems, presents a description and analysis of several algorithms, both exact and heuristic, for solving the discrete optimization problem, and presents a computational comparison of these algorithms for the 2006 Recommended Childhood Immunization Schedule and several randomly generated childhood immunization schedules that may be representative of future childhood immunization schedules.

Chapter 4 presents general models (formulated as a decision problem and as a discrete optimization problem) that determine the set of vaccines (i.e., a vaccine formulary) that should be used in a clinical environment to satisfy any given childhood immunization schedule while restricting extraimmunization, rigorously explores the theoretical structure of this general model, and provides an extensive computational study. Specifically, Chapter 4 presents the computational complexity of the decision/discrete optimization problems, presents a description and analysis of several algorithms, both exact and heuristic, for solving the discrete optimization problem, and presents a computational comparison of these algorithms for the 2006 Recommended Childhood Immunization Schedule and several randomly generated childhood immunization schedules that may be representative of future childhood immunization schedules.

Chapter 5 extends the models presented in Chapters 3 and 4 by presenting general models (formulated as a decision problem and as a discrete optimization problem) that determine the set of vaccines (i.e., a vaccine formulary) that should be used in a clinical environment to satisfy any given childhood immunization schedule at minimum cost while also restricting extraimmunization. Specifically, Chapter 5 presents the computational complexity of the decision/discrete optimization problems, presents several formulation extensions, and presents a description and analysis of several algorithms, both exact and heuristic, for solving the discrete optimization problem.

This dissertation concludes with Chapter 6, which presents a brief conclusion along with several research extensions.

Chapter 2: Literature Review

This chapter presents a literature review of earlier research where operations research techniques have been used to address pediatric immunization problems. Other relevant topics related to this research are also discussed, such as specific decision problems and discrete optimization problems that are foundational to the research contained in this dissertation.

2.1 Operations Research and Pediatric Immunization

This section reviews the operations research literature as it applies to pediatric immunization. Operations research techniques have been used to address pediatric immunization problems; however, most of the research to date addresses the economics of pediatric vaccine formulary design, combination vaccine pricing, and vaccine wastage (Jacobson et al. 2003b, 2004). Weniger et al. (1998) report the results of a pilot study that uses operations research methods to assess the economic value of vaccine formularies—the set of vaccines inventoried by an immunization clinic or pediatrician. Specifically, the Recommended Childhood Immunization Schedule for a subset of diseases (diphtheria, tetanus, pertussis, *Haemophilus influenzae* type b, and hepatitis B) and a reduced set of time periods (1mo, 2mo, 4mo, 6mo, 12-18mo, and 60mo) were modeled as an integer program (IP). The objective of this IP was to aide decision-makers in determining the vaccine formulary that minimized the cost to fully immunize a child against all five diseases. They describe how the model may be used to determine the 'best value' to vaccine purchasers and briefly describe how operations research models might help determine the economic value of new vaccines being researched and developed. Jacobson et al. (1999) present a more rigorous presentation of this pilot study and demonstrate how the model selects different vaccine formularies depending on the desired economic criteria. For example, the model was evaluated under the economic criterion: minimum total cost, maximum total cost, and minimum total cost with all manufacturers represented. McGuire (2003) performed a cost-effectiveness analysis to determine the price for new vaccines prior to the vaccine's development. This analysis showed that new vaccine prices should be considerably higher than prices that are currently paid for new vaccines.

Sewell et al. (2001) embed the IP from the pilot study into a bisection algorithm (Burden and Faires 1997) to “reverse engineer” the maximum inclusion prices (the maximum price at which a vaccine remains part of the optimal vaccine formulary) of four combination vaccines not yet licensed in the United States (at the time of publication). Sewell and Jacobson (2003) present a rigorous description of this study, including the complete IP model. This study shows how operations research can provide beneficial economic analysis to the pharmaceutical companies that develop and manufacture vaccines (see Jacobson et al. (2003a, 2005) for additional applications of this bisection algorithm). Jacobson and Sewell (2002) extended the bisection/IP algorithm approach by including it with Monte Carlo simulation, thereby determining a probability distribution for the price of the four potential combination vaccines. Finally, Jacobson et al. (2006) uses a stochastic inventory model to analyze the CDC-proposed vaccine stockpile levels to determine their adequacy. Given prespecified vaccination coverage rates, this analysis provides insight into what the pediatric vaccine stockpile levels should be and the amount of funding needed to achieve such levels.

2.2 Other Topics of Interest

This section presents other relevant topics related to this research such as specific decision problems and discrete optimization problems that are foundational to the research contained in this dissertation.

There are several decision problems and discrete optimization problems that are foundational to the models presented in this dissertation. Specifically, the problems described here are used for the complexity results in Chapters 3 and 4 and motivate the algorithms and heuristics described throughout this dissertation.

The first decision problem, Minimum Cover, is taken from Garey and Johnson (1979) and is known to be *NP*-complete. This problem is now formally stated.

Minimum Cover (MC)

Given: A collection C of subsets of a finite set S, and a positive integer $K \leq |C|$.

Question: Does C contain a cover of size K or less for set S (i.e., a subset $C' \subseteq C$ with $|C'| \leq K$ such that every element of S belongs to at least one member of C')?

In the literature, the corresponding optimization problem associated with Minimum Cover is the *Set-Covering problem*, where each subset in the collection C is assigned a weight (or cost) and a minimum weighted cover of S is sought (Nemhauser and Wolsey 1999). The Set-Covering problem is generally modeled as the following integer program

$$\text{Minimize} \qquad \sum_{j=1}^{n} c_j x_j$$

Subject to

$$\sum_{j=1}^{n} a_{ij} x_j \geq 1 \qquad \text{for all } i \in S$$

$$x_j \in \{0,1\} \qquad \text{for all } j \in C$$

where

- c_j is the weight (or cost) corresponding to subset $j \in C$
- a_{ij} = 1(0) if element $i \in S$ is contained (not contained) in subset $j \in C$
- x_j = 1(0) if, for subset $j \in C$, $j \in C'$ ($j \notin C'$) (i.e., $x_j = 1$ if subset $j \in C$ is contained in the cover C' for S, 0 otherwise).

If the constraint for each $i \in S$ is set to equality, then the problem is the *Set-Partitioning problem* (Nemhauser and Wolsey 1999). See Caprara et al. (2000) and Hoffman and Padberg (2006) for a survery of algorithms and heuristics for the Set-Covering problem.

Another decision problem used for the complexity results in Chapters 3 and 4 is 3DM, which is also taken from Garey and Johnson (1979) and is known to be *NP*-complete. This problem in now formally stated.

3-Dimensional Matching (3DM)

Given: Set $M \subseteq W \times Y \times Z$, where W, Y, and Z are disjoint sets each containing q elements. Therefore, $W = \{w_1,w_2,\ldots,w_q\}$, $Y = \{y_1,y_2,\ldots,y_q\}$, $Z = \{z_1,z_2,\ldots,z_q\}$, and $M = \{m_1,m_2,\ldots,m_k\}$ where $m_i = (w, y, z)$, $i = 1,2,\ldots,k$, such that $w \in W$, $y \in Y$, and $z \in Z$.

Question: Does M contain a matching (i.e., does there exist a subset $M' \subseteq M$ such that $|M'| = q$ and no two elements of M' agree in any coordinates)?

The next set of decision problems are variations of the Satisfiability problem (Garey and Johnson 1979). The first two problems (3-SAT and 1-in-3 3-SAT) are taken from Garey and Johnson (1979) and are known to be *NP*-complete. These problems are now formally stated.

3-Satisfiability (3-SAT)

Given: A set of n Boolean variables ($y_1,y_2,\ldots,y_n$) and a collection of m clauses over the n Boolean variables ($C_1,C_2,\ldots,C_m$) such that $|C_i| = 3$, $i = 1,2,\ldots,m$.

Question: Are there values for the n Boolean variables such that each clause has at least one true literal?

One-In-Three 3-Satisfiability (1-in-3 3-SAT)

Given: A set of n Boolean variables ($y_1,y_2,\ldots,y_n$) and a collection of m clauses over the n Boolean variables ($C_1,C_2,\ldots,C_m$) such that $|C_i| = 3$, $i = 1,2,\ldots,m$.

Question: Are there values for the n Boolean variables such that each clause has exactly one true literal?

The following decision problem is a variation of 1-in-3 3-SAT:

One-In-Three 3-Satisfiability with 2-Satisfiability (1-in-3 3-SAT with 2-SAT)

Given: A set of n Boolean variables ($y_1,y_2,\ldots,y_n$) and a collection of $m+n$ clauses over the n Boolean variables ($C_1,C_2,\ldots,C_{m+n}$) such that $|C_i| = 3$, $i = 1,2,\ldots,m$, and $|C_i| = 2$, $i = m+1,m+2,\ldots,m+n$ where clause $C_{m+k} = (y_k \vee (1 - y_k))$, $k = 1,2,\ldots,n$.

Question: Are there values for the n Boolean variables such that each clause has exactly one true literal?

Clearly, a transformation from 1-in-3 3-SAT to 1-in-3 3SAT with 2-SAT can be made in polynomial time. Furthermore, 1-in-3 3-SAT has a solution if an only if 1-in-3 3-SAT with 2-SAT has a solution, and hence, 1-in-3 3-SAT with 2-SAT is also *NP*-complete. A similar variation of 3-SAT to 3-SAT with 2-SAT is possible.

In the literature, the corresponding optimization problem associated with Satisfiability is the *MAX-SAT problem*, where a Boolean variable assignment that maximizes the number of satisfied clauses is sought (Hochbaum 1997). See Miltersen (2005) for a description of algorithms for Satisfiability and Asano and Williamson (2002) for heuristics and approximation algorithms for the MAX-SAT problem.

This dissertation uses principles and techniques from complexity theory (Garey and Johnson 1979), linear programming (Bazarra et al. 1990), dynamic programming and integer programming (Nemhauser and Wolsey 1999), approximation algorithms (Hochbaum 1997), and randomized algorithms (Motwani and Raghavan, 1995).

Chapter 3: The Vaccine Formulary Selection with Limited Budget Problem

This chapter extends the research described in Chapter 2 by generalizing the model for any given childhood immunization schedule and by rigorously exploring the theoretical structure of this general model. An extensive computational study is also presented. The chapter is organized as follows. Section 3.1 presents general models (formulated as a decision problem and as a discrete optimization problem) that determine the set of vaccines (i.e., a vaccine formulary) that should be used in a clinical environment to satisfy any given childhood immunization schedule, and presents the terminology that is used throughout the dissertation. Section 3.2 presents the computational complexity of the decision/discrete optimization problems. Section 3.3 presents a description and analysis of several algorithms, both exact and heuristic, for solving the discrete optimization problem. Section 3.4 presents a computational comparison of these algorithms for the 2006 Recommended Childhood Immunization Schedule and several randomly generated childhood immunization schedules that may be representative of future childhood immunization schedules.

3.1 Model Formulation and Terminology

This section presents a model formulation for a decision problem and a discrete optimization problem used to design a vaccine formulary that addresses the cost of satisfying a given childhood immunization schedule. Given a childhood immunization schedule, the decision problem, termed the Vaccine Formulary Selection with Limited Budget Problem (VFSLBP), asks whether it is possible to design a vaccine formulary within a specified budget. This problem is now formally stated.

Vaccine Formulary Selection with Limited Budget Problem (VFSLBP)

Given:

- A set of time periods, $T = \{1,2,\ldots,\tau\}$,
- a set of diseases, $D = \{1,2,\ldots,\delta\}$,
- a set of vaccines $V = \{1,2,\ldots,\upsilon\}$, available to be administered to immunize against the δ diseases,

- the number of doses of a vaccine that must be administered for immunization against the δ diseases, $n_1, n_2, \ldots, n_\delta$,
- the cost of each vaccine, $c_1, c_2, \ldots, c_\upsilon$,
- a budget B,
- a set of binary parameters that indicate which vaccines immunize against which diseases; therefore, $I_{vd} = 1$ if vaccine $v \in V$ immunizes against disease $d \in D$, 0 otherwise,
- a set of binary parameters that indicate the set of time periods in which a particular dose of a vaccine may be administered to immunize against a disease; therefore, $P_{djt} = 1$ if in time period $t \in T$, a vaccine may be administered to satisfy the j^{th} dose, $j = 1,2,\ldots,n_d$, requirement for disease $d \in D$, 0 otherwise,
- a set of binary parameters that indicate the set of time periods in which a vaccine may be administered to satisfy any dose requirement against a disease; therefore, $Q_{dt} = 1$ if in time period $t \in T$, a vaccine may be administered to satisfy any dose requirement against disease $d \in D$, 0 otherwise, (i.e., for any disease $d \in D$ and time period $t \in T$, $Q_{dt} = 1$ if and only if $P_{djt} = 1$ for some dose $j = 1,2,\ldots,n_d$.),
- a set of integer parameters that indicate the minimum number of doses of a vaccine required for disease $d \in D$ through time period $t \in T$; denoted m_{dt}.

Question: Does there exist a set of vaccines from V that can be administered over the time periods in T such that these vaccines immunize against all the diseases in D, at a total cost no greater than B (i.e., do there exist values for the binary variables X_{tv}, $t \in T$, $v \in V$, where $X_{tv} = 1$ if vaccine $v \in V$ is administered in time period $t \in T$, 0 otherwise, and for the binary variables U_{dt}, $d \in D$, $t \in T$, where $U_{dt} = 1$ if any vaccine $v \in V$ that immunizes against disease $d \in D$ is administered in time period $t \in T$, 0 otherwise, such that for all diseases $d \in D$, $\Sigma_{t\in T}\Sigma_{v\in V}\, P_{djt}\, X_{tv}\, I_{vd} \geq 1$ for dose $j = 1,2,\ldots,n_d$, $\Sigma_{t=1,2,\ldots,t'}\, Q_{dt}\, U_{dt} \geq m_{dt'}$ for all time periods $t' \in T$, and $\Sigma_{v\in V}\, Q_{dt}\, X_{tv}\, I_{vd} \geq U_{dt}$ for all time periods $t \in T$; and $\Sigma_{t\in T}\Sigma_{v\in V}\, c_v\, X_{tv} \leq B$)?

In the formulation of VFSLBP, the given sets and parameters equate to a childhood immunization schedule together with budget and vaccine cost information. Unless otherwise stated, the phrase "childhood immunization schedule" refers to an arbitrary general immunization schedule, whereas the phrase “Recommended Childhood Immunization Schedule” refers to the published CDC immunization schedule in Figure 1. Furthermore, the doses for all diseases $d \in D$ are assumed to be *sequentially ordered*, which means that for all doses j, $k = 1,2,\ldots,n_d$, $j < k$, there exists a time period $t' \in T$ such that $P_{djt'} = 1$ and $P_{dkt} = 0$ for all $t \leq t'$, $t \in T$. Define the *valency*, denoted by $Val(v)$, as the number of antigens contained in vaccine $v \in V$, and hence, $Val(v) = \Sigma_{d\in D} I_{vd}$. Combination vaccines are often referred to as *multivalent* vaccines, or simply *multivalents*, because $Val(v) \geq 2$ when $v \in V$ is a combination vaccine. Furthermore, vaccine $v \in V$, where $Val(v) = 1, 2, 3, 4, 5$, or 6 is often referred to as a monovalent, bivalent, trivalent, tetravalent, pentavalent, or hexavalent vaccine, respectively. In practice, the dose parameters, n_d and m_{dt}, and schedule parameters, P_{djt} and Q_{dt}, depend on biological constraints and are determined by the recommendations of the ACIP and AAFP (CDC 2002). The cost parameter, c_v, is a general parameter that quantifies the economic cost of vaccine $v \in V$. For example, Weniger et al. (1998) considered the actual vaccine purchase price, the cost of preparing the vaccine by medical staff, and the cost of administration (needle/syringe, needle-free injections, or oral) for a given vaccine $v \in V$. The question in VFSLBP asks if there exists a vaccine formulary administered over the time periods in T that *satisfies* a given childhood immunization schedule and is within the given budget B (i.e., a variable assignment for the binary variables X_{tv}, for all time periods $t \in T$ and vaccines $v \in V$, and U_{dt}, for all diseases $d \in D$ and time periods $t \in T$, that satisfies the per dose requirements ($\Sigma_{t\in T}\Sigma_{v\in V}\ P_{djt}\ X_{tv}\ I_{vd} \geq 1$ for dose $j = 1,2,\ldots,n_d$) and total dosage requirements ($\Sigma_{t=1,2,\ldots,t'}\ Q_{dt}\ U_{dt} \geq m_{dt'}$ for all time period $t' \in T$ and $\Sigma_{v\in V}\ Q_{dt}\ X_{tv}\ I_{vd} \geq U_{dt}$ for all time periods $t \in T$) for each disease $d \in D$, and the budget constraint ($\Sigma_{t\in T}\Sigma_{v\in V}\ c_v\ X_{tv} \leq B$)). Therefore, VFSLBP permits extraimmunization (i.e., multiple vaccinations for disease $d \in D$ in time period $t \in T$, or vaccinations for disease $d \in D$ in time periods $t \in T$ when $P_{djt} = 0$ for all doses $j = 1,2,\ldots,n_d$). Chapter 4 considers the case when extraimmunization is restricted.

This decision problem can be addressed by solving a discrete optimization problem. More specifically, the following binary integer program can be used to answer VFSLBP.

Integer Programming Model for Vaccine Formulary Selection with Limited Budget Problem (VFSLBP(O))

Minimize $\sum_{t\in T}\sum_{v\in V} c_v X_{tv}$ (O)

Subject to

$\sum_{t\in T}\sum_{v\in V} P_{djt} X_{tv} I_{vd} \geq 1$ for all $d \in D$, $j = 1,2,\ldots,n_d$ (1)

$\sum_{t=1,2,\ldots,t'} Q_{dt} U_{dt} \geq m_{dt'}$ for all $d \in D$, $t' \in T$, (2)

$\sum_{v\in V} Q_{dt} X_{tv} I_{vd} \geq U_{dt}$ for all $d \in D$, $t \in T$, (3)

$X_{tv} \in \{0,1\}$ for all $t \in T$, $v \in V$, (4)

$U_{dt} \in \{0,1\}$ for all $d \in D$, $t \in T$, (5)

where sets T, D, and V, parameters $\{c_v\}$, $\{n_d\}$, $\{P_{djt}\}$, $\{I_{vd}\}$, $\{Q_{dt}\}$, and $\{m_{dt}\}$, and variables $\{X_{tv}\}$ and $\{U_{dt}\}$ are defined in VFSLBP.

The objective function (O) minimizes the total cost of the vaccine formulary subject to the dose requirements for each disease $d \in D$. Therefore, if the minimum total cost is less than or equal to the specified budget B, then the answer to VFSLBP is "yes." Constraint (1) ensures that for each disease $d \in D$, at least one vaccine that provides immunization for disease $d \in D$ is administered in some time period when dose $j = 1,2,\ldots,n_d$ may be administered. Constraint (2) and (3) guarantees that for each disease $d \in D$, at least m_{dt} doses of a vaccine that immunize against disease $d \in D$ are administered in the first $t \in T$ time periods, while also ensuring that at most one dose requirement for disease $d \in D$ is satisfied in time period $t \in T$. Finally, constraints (4) and (5) are the binary requirements for each decision variable.

To simplify the formulation of VFSLBP(O), define $T_{dj} = \{t \in T : P_{djt} = 1\}$ to be the set of time periods when dose $j = 1,2,\ldots,n_d$, may be administered for disease $d \in D$. Unless otherwise stated, assume that for all diseases $d \in D$ and doses $j = 1,2,\ldots, n_d$, the time periods in T_{dj} are consecutive. A disease $d \in D$ is defined to have *mutually exclusive doses* if $T_{di} \cap T_{dj} = \emptyset$ for all i, $j = 1,2,\ldots,n_d$, $i \neq j$ (i.e., the sets T_{dj}, $j = 1,2,\ldots,n_d$ are pairwise mutually exclusive), and hence, the set of time periods when dose j may be administered for disease $d \in D$ does not overlap with the set of time periods when dose i may be administered for all $i \neq j$. For example, disease 1 in Figure 2 does not have mutually exclusive doses. Note that constraints (2) and (3) are redundant for any disease $d \in D$ with mutually exclusive doses.

Therefore, if every disease has mutually exclusive doses then VFSLBP(O) simplifies to the same objective function (O) with constraints (1) and (4) only. This simplified problem is denoted by VFSLBP(O)-MED to signify the optimization model where each disease $d \in D$ has mutually exclusive doses. All of the diseases in the 2006 Recommended Childhood Immunization Schedule have mutually exclusive doses, though some diseases in past schedules did not have this property. For example, hepatitis B did not have mutually exclusive doses in the 2005 Recommended Childhood Immunization Schedule (CDC 2005).

<table>
<tr><th rowspan="2">DISEASE</th><th colspan="8">TIME PERIOD</th></tr>
<tr><th>1</th><th>2</th><th>3</th><th>4</th><th>5</th><th>6</th><th>7</th><th>8</th></tr>
<tr><td rowspan="2">1</td><td colspan="3">Dose 1</td><td></td><td colspan="4">Dose 3</td></tr>
<tr><td></td><td colspan="3">Dose 2</td><td></td><td></td><td></td><td></td></tr>
<tr><td>2</td><td></td><td></td><td>Dose 1</td><td>Dose 2</td><td colspan="4">Dose 3</td></tr>
<tr><td>3</td><td></td><td></td><td></td><td></td><td></td><td colspan="3">Dose 1</td></tr>
</table>

Figure 2: Childhood Immunization Schedule without Mutually Exclusive Doses

Example 1

An example of the model parameters and formulations are now given for the childhood immunization schedule in Figure 2. Specifically, $T = \{1,2,3,4,5,6,7,8\}$, $D = \{1,2,3\}$, and the dose vector $n = (3,3,1)$, where the d^{th} component of n corresponds to the dose requirement for disease $d = 1,2,3$. Furthermore, the binary schedule parameters P_{djt} and Q_{dt} are:

for disease $d = 1$,

dose $j = 1$: $P_{djt} = 1(0)$ for time period $t = 1,2,3(4,5,6,7,8)$,

dose $j = 2$: $P_{djt} = 1(0)$ for time period $t = 2,3,4(1,5,6,7,8)$, and

dose $j = 3$; $P_{djt} = 1(0)$ for time period $t = 5,6,7,8(1,2,3,4)$, and hence,

$Q_{dt} = 1$ for all time periods $t = 1,2,\ldots,8$,

for disease $d = 2$,

dose $j = 1$: $P_{djt} = 1(0)$ for time period $t = 3(1,2,4,5,6,7,8)$,

dose $j = 2$: $P_{djt} = 1(0)$ for time period $t = 4(1,2,3,5,6,7,8)$, and

dose $j = 3$: $P_{djt} = 1(0)$ for time period $t = 5,6,7,8(1,2,3,4)$, and hence,

$Q_{dt} = 1(0)$ for time period $t = 3,4,5,6,7,8(1,2)$, and

for disease $d = 3$,

dose $j = 1$: $P_{djt} = 1(0)$ for time period $t = 6,7,8(1,2,3,4,5)$, and hence,

$Q_{dt} = 1(0)$ for time period $t = 6,7,8(1,2,3,4,5)$.

Finally, the minimum dose vectors for each disease $d \in D$ are m_1 = (0,0,1,2,2,2,2,3), m_2 = (0,0,1,2,2,2,2,3), and m_3 = (0,0,0,0,0,0,0,1), where m_{dt} is the t^{th} component, t = 1,2,…,8, of vector m_d for disease d = 1,2,3.

Suppose V = {1 = {1}, 2 = {2}, 3 = {3}, 4 = {1,2,3}}, which implies the binary parameters I_{vd}: I_{1d} = 1(0) for disease d = 1(2,3), I_{2d} = 1(0) for disease d = 2(1,3), I_{3d} = 1 for disease d = 3(1,2), and I_{4d} = 1 for all diseases d = 1,2,3. Therefore, each vaccine $v \in V$ may be interpreted as the subset of diseases from the set D for which the vaccine provides immunization against. Moreover, let c = (1,2,2,3) be the cost vector, where the v^{th} component of c corresponds to the cost of vaccine v = 1,2,3,4, and let B = 9. Therefore, VFSLBP asks: do there exist values for the binary variables X_{tv}, $t \in T$, $v \in V$, where X_{tv} = 1 if vaccine $v \in V$ is administered in time period $t \in T$, 0 otherwise, and for binary variables U_{dt}, $d \in D$, $t \in T$, where U_{dt} = 1 if any vaccine $v \in V$ that immunizes against disease $d \in D$ is administered in time period $t \in T$, 0 otherwise, such that:

for disease d = 1,

dose j = 1: $X_{11} + X_{14} + X_{21} + X_{24} + X_{31} + X_{34} \geq 1$,

dose j = 2: $X_{21} + X_{24} + X_{31} + X_{34} + X_{41} + X_{44} \geq 1$,

dose j = 3: $X_{51} + X_{54} + X_{61} + X_{64} + X_{71} + X_{74} + X_{81} + X_{84} \geq 1$, and time period

t = 1: $U_{11} \geq 0$,

t = 2: $U_{11} + U_{12} \geq 0$,

t = 3: $U_{11} + U_{12} + U_{13} \geq 1$,

t = 4: $U_{11} + U_{12} + U_{13} + U_{14} \geq 2$,

t = 5: $U_{11} + U_{12} + U_{13} + U_{14} + U_{15} \geq 2$,

t = 6: $U_{11} + U_{12} + U_{13} + U_{14} + U_{15} + U_{16} \geq 2$,

t = 7: $U_{11} + U_{12} + U_{13} + U_{14} + U_{15} + U_{16} + U_{17} \geq 2$,

t = 8: $U_{11} + U_{12} + U_{13} + U_{14} + U_{15} + U_{16} + U_{17} + U_{18} \geq 3$,

t = 1: $X_{11} + X_{14} \geq U_{11}$,

t = 2: $X_{21} + X_{24} \geq U_{12}$,

t = 3: $X_{31} + X_{34} \geq U_{13}$,

t = 4: $X_{41} + X_{44} \geq U_{14}$,

t = 5: $X_{51} + X_{54} \geq U_{15}$,

t = 6: $X_{61} + X_{64} \geq U_{16}$,

t = 7: $X_{71} + X_{74} \geq U_{17}$,

t = 8: $X_{81} + X_{84} \geq U_{18}$,

for disease d = 2,

dose j = 1: $X_{32} + X_{34} \geq 1$,

dose j = 2: $X_{42} + X_{44} \geq 1$,

dose j = 3: $X_{52} + X_{54} + X_{62} + X_{64} + X_{72} + X_{74} + X_{82} + X_{84} \geq 1$, and time period

t = 1: $Q_{21} = 0 \Rightarrow$ no constraint,

t = 2: $Q_{22} = 0 \Rightarrow$ no constraint,

t = 3: $U_{23} \geq 1$,

t = 4: $U_{23} + U_{24} \geq 2$,

t = 5: $U_{23} + U_{24} + U_{25} \geq 2$,

t = 6: $U_{23} + U_{24} + U_{25} + U_{26} \geq 2$,

t = 7: $U_{23} + U_{24} + U_{25} + U_{26} + U_{27} \geq 2$,

t = 8: $U_{23} + U_{24} + U_{25} + U_{26} + U_{27} + U_{28} \geq 3$,

t = 1: $Q_{21} = 0 \Rightarrow$ no constraint,

t = 2: $Q_{22} = 0 \Rightarrow$ no constraint,

t = 3: $X_{32} + X_{34} \geq U_{23}$,

t = 4: $X_{42} + X_{44} \geq U_{24}$,

t = 5: $X_{52} + X_{54} \geq U_{25}$,

t = 6: $X_{62} + X_{64} \geq U_{26}$,

t = 7: $X_{72} + X_{74} \geq U_{27}$,

t = 8: $X_{82} + X_{84} \geq U_{28}$, and

for disease d = 3,

dose j = 1: $X_{63} + X_{64} + X_{73} + X_{74} + X_{83} + X_{84} \geq 1$, and time period

t = 1: $Q_{31} = 0 \Rightarrow$ no constraint,

t = 2: $Q_{32} = 0 \Rightarrow$ no constraint,

t = 3: $Q_{33} = 0 \Rightarrow$ no constraint,

t = 4: $Q_{34} = 0 \Rightarrow$ no constraint,

t = 5: $Q_{35} = 0 \Rightarrow$ no constraint,

t = 6: $U_{36} \geq 0$,

t = 7: $U_{36} + U_{37} \geq 0$,

t = 8: $U_{36} + U_{37} + U_{38} \geq 1$,

t = 1: $Q_{31} = 0 \Rightarrow$ no constraint,

t = 2: $Q_{32} = 0 \Rightarrow$ no constraint,

t = 3: $Q_{33} = 0 \Rightarrow$ no constraint,

t = 4: $Q_{34} = 0 \Rightarrow$ no constraint,

t = 5: $Q_{35} = 0 \Rightarrow$ no constraint,

t = 6: $X_{63} + X_{64} \geq U_{36}$,

t = 7: $X_{73} + X_{74} \geq U_{37}$,

t = 8: $X_{83} + X_{84} \geq U_{38}$, and

for budget B,

$$\sum_{t=1}^{8} X_{t1} + 2\sum_{t=1}^{8} X_{t2} + 2\sum_{t=1}^{8} X_{t3} + 3\sum_{t=1}^{8} X_{t4} \leq 9\,?$$

For diseases d = 2,3, note that the constraints $\Sigma_{t=1,2,\ldots,t'}\ Q_{dt}\ U_{dt} \geq m_{dt'}$ for time periods t' $\in T$ and $\Sigma_{v\in V}\ Q_{dt}\ X_{tv}\ I_{vd} \geq U_{dt}$ for time periods $t \in T$ are redundant, since both diseases have mutually exclusive doses. Therefore, the formulation for VFSLBP(O) (excluding redundant constraints) for this example is:

Minimize $\quad \sum_{t=1}^{8} X_{t1} + 2\sum_{t=1}^{8} X_{t2} + 2\sum_{t=1}^{8} X_{t3} + 3\sum_{t=1}^{8} X_{t4}$

Subject to

$$X_{11} + X_{14} + X_{21} + X_{24} + X_{31} + X_{34} \geq 1$$

$$X_{21} + X_{24} + X_{31} + X_{34} + X_{41} + X_{44} \geq 1$$

$$X_{51} + X_{54} + X_{61} + X_{64} + X_{71} + X_{74} + X_{81} + X_{84} \geq 1$$

$$U_{11} + U_{12} + U_{13} \geq 1$$

$$U_{11} + U_{12} + U_{13} + U_{14} \geq 2$$

$$U_{11} + U_{12} + U_{13} + U_{14} + U_{15} + U_{16} + U_{17} + U_{18} \geq 3$$

$$X_{11} + X_{14} \geq U_{11}$$

$$X_{21} + X_{24} \geq U_{12}$$

$$X_{31} + X_{34} \geq U_{13}$$

$$X_{41} + X_{44} \geq U_{14}$$

$$X_{51} + X_{54} \geq U_{15}$$

$$X_{61} + X_{64} \geq U_{16}$$

$$X_{71} + X_{74} \geq U_{17}$$

$$X_{81} + X_{84} \geq U_{18}$$

$$X_{32} + X_{34} \geq 1$$

$$X_{42} + X_{44} \geq 1$$

$$X_{52} + X_{54} + X_{62} + X_{64} + X_{72} + X_{74} + X_{82} + X_{84} \geq 1$$

$$X_{63} + X_{64} + X_{73} + X_{74} + X_{83} + X_{84} \geq 1$$

$$X_{tv} \in \{0,1\} \qquad \text{for all } t \in T,\ v \in V$$

$$U_{1t} \in \{0,1\} \qquad \text{for all } t \in T. \ \square$$

3.2 Computational Complexity

This section presents the computational complexity of VFSLBP and VFSLBP(O). Not surprisingly, in the worst case, these problems are shown to be intractable. There are, however, some special cases that are solvable in polynomial time. Theorem 1 states that VFSLBP is *NP*-complete.

THEOREM 1: *VFSLBP is NP-complete in the strong sense.*

PROOF: First, VFSLBP is in the class *NP* since given a set of guessed values for the binary variables X_{tv}, $t \in T$, $v \in V$, and U_{dt}, $d \in D$, $t \in T$, $\Sigma_{t\in T}\Sigma_{v\in V} P_{djt} X_{tv} I_{vd} \geq 1$ for all $d \in D$, $j = 1,2,\ldots,n_d$, $\Sigma_{t=1,2,\ldots,t'} Q_{dt} U_{dt} \geq m_{dt'}$ for all $d \in D$ and $t' \in T$, and $\Sigma_{v\in V} Q_{dt} X_{tv} I_{vd} \geq U_{dt}$ for all $d \in D$ and $t \in T$ can all be checked in $O(\upsilon\tau^2\delta)$ time, and $\Sigma_{t\in T}\Sigma_{v\in V} c_v X_{tv} \leq B$ can be checked in $O(\upsilon\tau)$ time.

To complete the proof, a polynomial transformation from Minimum Cover (MC) to VFSLBP is constructed.

Given an arbitrary instance of MC, define a particular instance of VFSLBP as follows: Set $T = \{1\}$, $D = S$, $V = C$, $n_1 = n_2 = \ldots = n_\delta = 1$, $c_1 = c_2 = \ldots = c_\upsilon = 1$, and $B = K$. If $d \in S$ is in subset $v \in C$, set the binary parameters $I_{vd} = 1$ and 0 otherwise. Lastly, $P_{djt} = 1$, $d \in D$, $j = 1$, $t = 1$, $Q_{dt} = 1$, and $m_{dt} = 1$ for $d \in D$, $t = 1$. Clearly, this transformation can be done in polynomial time in the size of the arbitrary instance of MC.

It remains to show that a *yes* for the particular instance of VFSLBP implies a *yes* for the arbitrary instance MC, and a *yes* for the arbitrary instance MC implies a *yes* for the particular instance of VFSLBP.

Suppose the answer to the particular instance of VFSLBP is *yes*. Then there exist values for the binary variables X_{1v}, $v \in V$, such that $\Sigma_{v \in V} X_{1v} I_{vd} \geq 1$ for all $d \in D$, and $\Sigma_{v \in V} X_{1v} \leq B$. (The binary variables U_{dt}, $d \in D$, $t = 1$ may be ignored, since $U_{d1} \geq 1$ for all $d \in D$.) The claim is that if $X_{1v} = 1$, then the corresponding subset $v \in C$ is part of a cover $C' \subseteq C$ for S (of size K or less). Let C' be the set of subsets of C corresponding to $X_{1v} = 1$. The first constraint ensures that for $d \in S$, there exists at least one subset C' such that $d \in C'$. The second constraint ensures that $|C'| \leq K$. Therefore, the answer to the arbitrary instance of MC is *yes*.

Now suppose the answer to the arbitrary instance of MC is *yes*. Then there exists a set of K or fewer subsets $C' \subseteq C$ such that for all $d \in S$, d is in at least one element of C'. Since each subset in C for the arbitrary instance of MC corresponds to a different vaccine for the particular instance of VFSLBP, set $X_{1v} = 1$ if the corresponding subset is in C', and zero otherwise. Moreover, each element in S for the arbitrary instance of MC corresponds to a different disease for the particular instance of VFSLBP. Therefore, for each $d \in D$, $\Sigma_{v \in V} X_{1v} I_{vd}$ computes the number of subsets in C' that cover $d \in S$ (for the MC instance), and hence, if C' is a cover for S, then $\Sigma_{v \in V} X_{1v} I_{vd} \geq 1$ for all $d \in D$. Furthermore, since $|C'| \leq K$, then $\Sigma_{v \in V} X_{1v} \leq B$. Therefore, the answer to the particular instance of VFSLBP is *yes*.

By definition, a problem **P** is *NP-complete in the strong sense* if there exists a polynomial p such that $\mathbf{P}_p$ is *NP-complete*, where $\mathbf{P}_p$ is the set of all instances **I** of **P**, such that Max[**I**] $\leq p$(Length[**I**]), where Max[**I**] represents the magnitude of the largest number occurring in the instance **I** and Length[**I**] represents the length of a "reasonable and concise" encoding of instance **I** (Garey and Johnson 1979). This definition implies that VFSLBP is *NP-complete in the strong sense* if VFSLBP may be restricted so that Max[**I**] $\leq M$ for some constant M and the restricted problem remains *NP-complete*. Given an instance **I** of

VFSLBP, let Max[**I**] = $\tau(\max_{d\in D} n_d)(\max_{v\in V} c_v)$. When VFSLBP is restricted by letting τ = 1, $n_1 = n_2 = \ldots = n_\delta = 1$, and $c_1 = c_2 = \ldots = c_\upsilon = 1$ as described in the transformation above, the problem becomes MC. Therefore, Max[**I**] is bounded by a constant and the restricted VFSLBP is the *NP-complete* problem MC, and hence, VFSLBP is *NP-complete in the strong sense*. ■

The proof of Theorem 1 suggests several special cases of VFSLBP that remain *NP*-complete. Specifically, VFSLBP remains *NP*-complete when there exists only one time period (i.e., $\tau = 1$), when the vaccine costs are equal (i.e., $c_v = c$ for all vaccines $v \in V$), and when each disease requires only one dose of vaccine (i.e., $n_d = 1$ for all diseases $d \in D$). Theorem 2 gives some additional special cases of VFSLBP that remain *NP*-complete.

THEOREM 2: *The following special cases of VFSLBP are NP-complete*:

i) *Only one vaccine exists* (*i.e.*, $\upsilon = 1$ *where* $I_{\upsilon d} = 1$ *for all diseases* $d \in D$),

ii) *The disease set has cardinality of at least three* (*i.e.*, $\delta \geq 3$),

iii) *Every vaccine is at least a trivalent vaccine* (*i.e.*, $Val(v) \geq 3$ *for all vaccines* $v \in V$).

PROOF: To show *i*), an alternative polynomial transformation from MC to VFSLBP is constructed, where $V = \{1\}$.

Given an arbitrary instance of MC, define a particular instance of VFSLBP as follows: Set $T = C$, $D = S$, $V = \{1\}$, $n_1 = n_2 = \ldots = n_\delta = 1$, $c_1 = 1$, and $B = K$. Set the binary parameters $I_{vd} = 1$ for $v = 1$, $d \in D$. Set, $P_{djt} = 1$ for $d \in D$, $j = 1$, and $t \in T$ if $d \in S$ is in subset $t \in C$, and 0 otherwise. Lastly, $Q_{dt} = m_{dt} = 1$ whenever $P_{d1t} = 1$, and 0 otherwise. Clearly, this transformation is polynomial in the size of the arbitrary instance of MC.

Suppose the answer to the particular instance of VFSLBP is *yes*. Then there exist values for the binary variables X_{t1}, $t \in T$, such that $\Sigma_{t\in T} P_{djt} X_{t1} \geq 1$ for all $d \in D$, $j = 1$, and $\Sigma_{t\in T} X_{t1} \leq B$. (The constraints $\Sigma_{t\in T} P_{djt} X_{t1} \geq 1$ for all $d \in D$, $j = 1$ imply the constraints with binary variables U_{dt}, $d \in D$, $t \in T$, $\Sigma_{t'=1,2,\ldots,t'} Q_{dt} U_{dt} \leq m_{dt'}$ for all $d \in D$, $t' \in T$, and $X_{t1} \geq U_{dt}$ are also satisfied.) The claim is that if $X_{t1} = 1$, then the corresponding subset $t \in C$ is part of a cover $C' \subseteq C$ for S (of size K or less). Let C' be the set of subsets of C corresponding to $X_{t1} = 1$. The first constraint ensures that for $d \in S$, there exists at least one subset C' such that $d \in C'$. The second constraint ensures that $|C'| \leq K$. Therefore, the answer to the arbitrary instance of MC is *yes*.

Now suppose the answer to the arbitrary instance of MC is *yes*. Then there exists a set of K or fewer subsets $C' \subseteq C$ such that for all $d \in S$, d is in at least one element of C'. Since each subset in C for the arbitrary instance of MC corresponds to a different time period for the particular instance of VFSLBP, set $X_{t1} = 1$ if the corresponding subset is in C', and zero otherwise. Moreover, each element in S for the arbitrary instance of MC corresponds to a different disease for the particular instance of VFSLBP. Therefore, for each $d \in D$, $\Sigma_{t\in T} P_{d1t} X_{t1}$ computes the number of subsets in C' that cover $d \in S$ (for the MC instance). Thus if C' is a cover for S, then $\Sigma_{t\in T} P_{d1t} X_{t1} = \Sigma_{t\in T} Q_{dt} X_{t1} \geq 1$ for all $d \in D$. Furthermore, since $|C'| \leq K$, then $\Sigma_{t\in T} X_{t1} \leq B$. Therefore, the answer to the particular instance of VFSLBP is *yes*.

To show *ii*) and *iii*), a polynomial transformation from 3DM to VFSLBP is constructed, where $\delta = 3$ and $V = \{1\}$ is a trivalent vaccine.

Given an arbitrary instance of 3DM define a particular instance of VFSLBP as follows: set $T = M$, $D = \{1,2,3\}$, $V = \{1\}$, $n_1 = n_2 = n_3 = q$, $c_1 = 1$, and $B = q$. Let the q elements in W, Y, and Z correspond to the doses of disease 1, 2, and 3, respectively. Hence, w_1 corresponds to the first dose of vaccine for disease 1, w_2 corresponds to the second dose of vaccine for disease 1, and so forth through dose q. Furthermore, since $T = M$, then the 3-tuple m_i, $i = 1,2,...,k$, corresponds to the i^{th} time period. Set the binary parameters $I_{1d} = 1$ for all $d \in D$. Set $P_{1jt} = 1$, $j = 1,2,\ldots q$, $t = i = 1,2,\ldots,k$, if element $w_j \in m_i$, 0 otherwise; likewise, $P_{2jt} = 1$, $j = 1,2,\ldots q$, $t = i = 1,2,\ldots,k$, if element $y_j \in m_i$, 0 otherwise; and, finally, $P_{3jt} = 1$, $j = 1,2,\ldots q$, $t = i = 1,2,\ldots,k$, if element $z_j \in m_i$, 0 otherwise. Lastly, set $Q_{dt} = 1$ for all $d \in D$ and $t \in T$ since in every time period some dose of vaccine v is permitted for disease d, and set $m_{dt} = 0$ for all $d \in D$, $t = 1,2,\ldots,k-1$, and $m_{dt} = q$ for $d \in D$, $t = k$. Clearly, this transformation is polynomial in the size of the arbitrary instance of 3DM.

Suppose the answer to the particular instance of VFSLBP is *yes*. Then there exist values for the binary variables X_{t1}, $t = 1,2,\ldots,|M|$, and U_{dt}, $d \in D$, $t \in T$, such that

$$\sum_{t\in T} P_{djt} X_{t1} \geq 1 \quad \text{for all } d \in D, j = 1,2,..,q; \tag{1}$$

$$\sum_{t\in T} Q_{dt} U_{dt} \geq q \quad \text{for all } d \in D; \tag{2}$$

$$X_{t1} \geq U_{dt} \qquad \text{for all } d \in D, t \in T; \tag{3}$$

$$\sum_{t\in T} \sum_{v\in V} c_v X_{t1} = \sum_{t\in T} X_{t1} \leq q \text{ (since } c_1 = 1). \tag{4}$$

Observe that in any time period t, exactly one dose for each disease $d \in D$ is permitted, that is, given d and t, $P_{djt} = 1$ for some $j = 1,2,\ldots,q$. Therefore, constraints (2) and (3) imply $\Sigma_{t\in T} Q_{dt} X_{t1} \geq q$ for all $d \in D$ and becomes the single constraint $\Sigma_{t\in T} X_{t1} \geq q$. This observation along with constraint (4) (i.e., $\Sigma_{t\in T} X_{t1} \leq q$) implies $\Sigma_{t\in T} X_{t1} = q$, which also means constraints (1) are tight (i.e., $\Sigma_{t\in T} P_{djt} X_{t1} = 1$ for all $d \in D$, $j = 1,2,..,q$) or some dose would not be satisfied. Hence, the constraints for this particular instance of VFSLBP become

$$\sum\nolimits_{t\in T} P_{djt} X_{t1} = 1 \quad \text{for all } d \in D, j = 1,2,..,q; \tag{1'}$$

$$\sum\nolimits_{t\in T} X_{t1} = q. \tag{2'}$$

The claim is that if $X_{t1} = 1$, then the 3-tuple $m \in M$ corresponding to time period t is part of the matching $M' \subseteq M$. Let M' be the set of 3-tuples in M corresponding to $X_{t1} = 1$. Constraints (1') ensure there exists exactly one time period $t \in T$ that vaccine 1 is administered to satisfy each dose requirement for every $d \in D$. Since the j^{th} dose corresponds to some element in $w_j \in W$, $y_j \in Y$, or $z_j \in Z$, then exactly one $m \in M'$ contains element w_j, y_j, and z_j. This means that no two elements of M' agree in any coordinate. Constraint (2') ensures that vaccine 1 is administered in exactly q time periods, and hence, $|M'| = q$. Therefore, the answer to the arbitrary instance of 3DM is *yes*.

Now suppose the answer to the arbitrary instance of 3DM is *yes*. Then there exists a matching $M' \subseteq M$ such that $|M'| = q$ and no two elements of M' agree in any coordinate. Since each 3-tuple in M for the arbitrary instance of 3DM corresponds to a different time period for the particular instance of VFSLBP, set $X_{t1} = 1$ if the corresponding 3-tuple m is in M', and zero otherwise. Since M' is a matching, then each element of W, Y, and Z exists in exactly one $m \in M'$, and hence, each dose for each disease is satisfied exactly once in the time period corresponding to m (i.e., $\Sigma_{t\in T} P_{djt} X_{t1} = 1$ for all $d \in D$, $j = 1,2,\ldots,q$). Moreover, $|M'| = q$, where no two elements of M' agree in any coordinate implies that $\Sigma_{t\in T} Q_{dt} X_{t1} = q$ for all $d \in D$, and hence, $\Sigma_{t\in T} P_{djt} X_{t1} \geq 1$ for all $d \in D$, $j = 1,2,\ldots,q$; $\Sigma_{t\in T} Q_{dt} X_{t1} \geq q$ for all $d \in D$; and $\Sigma_{t\in T} X_{t1} \leq q$. Therefore, these values of $X_{t1} = 1$ provide a *yes* answer to the particular instance of VFSLBP. ■

Theorems 1 and 2 imply VFSLBP remains *NP*-complete even when the sets T, D, and V, or when the dose (n_d, $d \in D$) and cost (c_v, $v \in V$) parameters are significantly restricted. In addition, since VFSLBP is *NP*-complete, then the corresponding optimization problem

VFSLBP(O) is *NP*-hard. Another facet to the complexity of VFSLBP lies in the flexibility of the childhood immunization schedule. In general, VFSLBP becomes more difficult if the doses for each disease may be administered in several time periods (i.e., for a given disease $d \in D$ and dose $j = 1,2,\ldots,n_d$, $P_{djt} = 1$ for multiple time periods $t \in T$). Define a childhood immunization schedule as *tight* if every required dose of vaccine for each disease $d \in D$ may be administered in exactly one time period (i.e., for dose $j = 1,2,\ldots,n_d$ and disease $d \in D$, $P_{djt} = 1$ for exactly one time period $t \in T$). By definition, a tight schedule is less flexible. A tight schedule also implies that all diseases $d \in D$ have mutually exclusive doses, since dose $j = 1,2,\ldots,n_d$ may be administered in exactly one time period, and hence, the time period $t \in T$ when $P_{djt} = 1$ is unique.

Special cases of VFSLBP(O) that are solvable in polynomial time occur when the valency of the vaccine set is limited to monovalent and bivalent vaccines, the schedule is tight, and when the number of diseases is less than three. To see this, first consider limitations on the valency of the vaccine set. Lemma 1 considers the case when all vaccines $v \in V$ are monovalents.

LEMMA 1: *If $Val(v) = 1$ for all vaccines $v \in V$, then VFSLBP(O) is solvable in $O(\tau \cdot \delta + \upsilon)$ time and has a minimum total cost of $\Sigma_{d \in D}\, c_d\, n_d$, where $c_d = \min\{c_v : I_{vd} = 1, v \in V\}$ for each disease $d \in D$.*

PROOF: Each disease $d \in D$ requires n_d doses at cost $c_d = \min\{c_v : I_{vd} = 1, v \in V\}$, which is the minimum cost of a vaccine $v \in V$ that immunizes against disease $d \in D$. Therefore, the minimum cost to satisfy a childhood immunization schedule is given by $\Sigma_{d \in D}\, c_d\, n_d$, and the optimal vaccine schedule may be found by looping through the set of time periods and diseases and administering dose $j = 1,2, \ldots,n_d$, in the first time period when $P_{djt} = 1$. ■

Define the linear programming (LP) relaxation of VFSLBP(O)-MED as the LP with objective function (O) subject to constraints (1) in VFSLBP(O)-MED and with the relaxed variable constraint $0 \leq X_{tv} \leq 1$ for all time periods $t \in T$ and vaccines $v \in V$. Theorem 3 states a stronger result than Lemma 1 for VFSLBP(O)-MED when all vaccines $v \in V$ are monovalents.

THEOREM 3: *If $Val(v) = 1$ for all vaccines $v \in V$, then the LP relaxation of VFSLBP(O)-MED yields a binary optimal solution.*

PROOF: This results assumes that the time periods where $P_{djt} = 1$ for all diseases $d \in D$, and doses $j = 1,2,\ldots,n_d$, are consecutive. This assumption holds in the 2006 Recommended Childhood Immunization Schedule, as well as all the randomly generated childhood immunization schedules presented in Section 3.4.

Suppose for a given instance of VFSLBP(O)-MED that $Val(v) = 1$ for all vaccines $v \in V$. Consider the LP relaxation of VFSLBP(O)-MED and denote its constraint matrix by **A**. If **A** is totally unimodular, then every basic feasible solution is integer, provided the right-hand-side (rhs) vector is integer (Ahuja et al. 1993). Clearly, the rhs vector in the LP relaxation of VFSLBP(O) is integer, and hence, it remains to show that **A** is indeed totally unimodular.

By definition, **A** is totally unimodular if every square submatrix of **A** has determinant 0, 1, or -1. It is well known that **A** is totally unimodular if the non-zero elements in each row are in consecutive columns (known as the *consecutive ones* property). Without loss of generality, assume that there is exactly one vaccine $v \in V$ that provides immunization against each disease $d \in D$. Furthermore, assume the columns of **A** are ordered according to the set V. For example, the first τ columns of **A** correspond to the decision variables associated with the first vaccine, (i.e., column t corresponds to decision variable X_{t1}), the second τ columns of **A** correspond to the decision variables associated with the second vaccine, and so on.

Now consider some disease $d \in D$, and let $v \in V$ be a vaccine such that $I_{vd} = 1$. Since $Val(v) = 1$, then the only non-zero entries in any row of **A** that corresponds to disease $d \in D$ must be in the τ consecutive columns corresponding to vaccine $v \in V$. By assumption, the time periods when $P_{djt} = 1$ for disease $d \in D$, dose $j = 1,2,\ldots,n_d$, are consecutive, and hence, the rows corresponding to each dose for disease $d \in D$ have the consecutive ones property, which implies that **A** is totally unimodular. ■

Theorem 3 also implies that VFSLBP(O)-MED is solvable in polynomial time when all vaccines $v \in V$ are monovalents, since LP is solvable in polynomial time (Bazaraa et al. 1990). Moreover, Theorem 3 may be used to show that the heuristics presented in Section 3.3 return the optimal solution when all vaccines $v \in V$ are monovalents.

Given a tight childhood immunization schedule and a vaccine set composed of monovalent and bivalent vaccines, Lemma 2 yields a second polynomial time solvable special case of VFSLBP(O).

LEMMA 2: *Given a tight childhood immunization schedule, if Val(v)* ≤ 2 *for all vaccines v* $\in$ *V, then VFSLBP(O) is solvable in* $O(\tau \cdot \delta^2)$ *time.*

PROOF: Consider some time period $t \in T$. Since the childhood immunization schedule is tight, if dose j for disease $d \in D$ may be administered (i.e., $P_{djt} = 1$), then it must be administered in time period $t \in T$. Let $D_t = \{d \in D: P_{djt} = 1 \text{ for some } j = 1,2,\ldots,n_d\}$ and $V_t = \{v \in V: I_{vd} = 1 \text{ and } d \in D_t\}$ for $t \in T$. Therefore, time period $t \in T$ yields a Set-Covering problem instance with base set D_t and a weighted collection of subsets V_t with weights c_v for each $v \in V_t$. Since the Set-Covering problem is solvable in polynomial time when each subset in the collection of subsets has cardinality of at most two, then finding the optimal vaccine set in time period $t \in T$ is solvable in polynomial time if each vaccine provides immunization for at most two diseases (i.e., monovalent and bivalent vaccines). Applying this result for all time periods $t \in T$ equates to at most τ Set-Covering problem instances, all of which are solvable in polynomial time using a matching algorithm with $O(\delta^2)$ complexity (Garey and Johnson 1979). Therefore, the overall complexity for this special case is $O(\tau \cdot \delta^2)$.
■

Lemmas 1 and 2 and Theorem 3 imply VFSLBP(O) is polynomial time solvable if the valency of the vaccine set is restricted and if the childhood immunization schedule is tight. Additionally, when $\delta = 1$, VFSLBP(O) is polynomial time solvable (i.e., $O(\tau)$ time) since it is a special case of Lemma 1. VFSLBP(O) is also polynomial time solvable when $\delta = 2$ using the dynamic programming algorithm presented in Section 3.3.1, since the subproblem solved at each stage of the dynamic program is a Set-Covering problem instance, which is polynomial time solvable when all vaccines are bivalents (Garey and Johnson 1979). The complexity of VFSLBP(O) with a general disease set D, where all vaccines are bivalents, remains an open question. However, a general instance of VFSLBP(O)-MED with bivalent vaccines is polynomial time solvable by transforming VFSLBP(O)-MED into a Set-Covering problem instance where each subset in the collection of subsets has cardinality of at most two.

Table 1 summarizes the complexity results for VFSLBP, with the following definitions corresponding to the listed parameters:

$\boldsymbol{n}$ = the dosage requirement for each disease $d \in D$ (i.e., $n_d = \boldsymbol{n}$ for all $d \in D$),

$\boldsymbol{c}$ = the cost for each vaccine $v \in V$ (i.e., $c_v = \boldsymbol{c}$ for all $v \in V$).

Table 1: Summary of Complexity Results for VFSLBP

	Time Periods	Diseases	Vaccines	# of Doses for each Disease	Cost of Vaccines
	S* = *T	***S* = *D***	***S* = *V***	***n* = |*S*|**	***c* = |*S*|**
***S* = ∅**	Undefined	Undefined	Infeasible	Polynomial	Polynomial
|*S*| = 1	*NP*-hard	Polynomial	*NP*-hard	*NP*-hard	*NP*-hard
|*S*| = 2	⇓	Polynomial	⇓	⇓	⇓
|*S*| ≥ 3		*NP*-hard			

3.3 Algorithms and Heuristics

Given that VFSLBP(O) is *NP*-hard, even when significantly restricting the cardinality of the input sets, it is likely that a significant amount of computing effort will be needed to find the optimal vaccine formulary for a given childhood immunization schedule Exact algorithms that guarantee optimality do exist for VFSLBP(O), but unless $P = NP$, these algorithms will always have a worst case complexity that is exponential in the size of the inputs. Therefore, it is useful (even necessary) to design heuristics that do not guarantee optimality but execute in time that is polynomial in the size of the inputs. This section discusses both exact algorithms and heuristics for VFSLBP(O). Section 3.3.1 presents an exact dynamic programming algorithm for VFSLBP(O). Section 3.3.2 presents two rounding heuristics (*Rounding* and *MAX Rounding*) for VFSLBP(O)-MED. Section 3.3.3 presents a *Primal-Dual* heuristic for VFSLBP(O)-MED. Section 3.3.4 presents a *Greedy* heuristic for VFSLBP(O)-MED. Lastly, Section 3.3.5 presents a *MAX Rounding*, *Primal-Dual*, and *Greedy* heuristic for VFSLBP(O). The heuristics for VFSLBP(O)-MED are shown to be approximation algorithms, which provide an approximation bound on the cost of the heuristic solution.

3.3.1 Dynamic Programming Algorithm

In Section 3.1, VFSLBP(O) is modeled as a binary integer programming (IP) problem, and hence, may be solved using several well known integer optimization techniques (such as branch and bound; see Nemhauser and Wolsey 1999). Another useful exact technique is dynamic programming (DP). Sewell et al. (2005) formulate a specific DP algorithm to find the minimal cost vaccine formulary for the 2005 Recommended Childhood Immunization Schedule. This section presents and analyzes a generalized DP algorithm for VFSLBP(O).

Given the stated set of inputs for VFSLBP(O) (i.e., set of time periods T, set of diseases D, set of vaccines V, required number of doses n_d for each $d \in D$, vaccine costs c_v for each $v \in V$, and binary parameters P, Q, and I), the DP algorithm solves VFSLBP(O) one period at a time beginning at the first time period (i.e., $t = 1$), and steps through each time period in T until $t = \tau$. Therefore, the set T defines the stages of the DP algorithm. In addition to the minimum dose parameter m_{dt}, $d \in D$, $t \in T$, define M_{dt} as the maximum number of doses of a vaccine required for disease $d \in D$ through time period $t \in T$.

Define a state in the DP algorithm as the number of doses of a vaccine that have been administered for each disease through time period $t \in T$. Formally, a state in time period $t \in T$ is a δ-dimensional vector $\mathbf{S}_t = (S_{t1}, S_{t2}, \ldots, S_{t\delta})$, where S_{td} is the number of doses of a vaccine that have been administered for disease $d = 1, 2, \ldots, \delta$, in time periods $1, 2, \ldots, t$. Therefore, the state space in time period $t \in T$ is $\Omega_t = \{\mathbf{S}_t \in \mathbf{Z}^{\delta} : m_{dt} \leq S_{td} \leq M_{dt}$ for all $d \in D\}$, where $\mathbf{Z}$ denotes the set of all integers. The decision in time period $t \in T$ is which vaccines to administer that immunize against the diseases requiring vaccination in this time period (i.e., the binary decision variables X_{tv}), and is represented by the δ-dimensional binary vector $\mathbf{Y}_t = (Y_{t1}, Y_{t2}, \ldots, Y_{t\delta})$, where $Y_{td} = 1$ implies $X_{tv} = 1$ for some vaccine $v \in V$ that immunizes against disease $d \in D$ (i.e., $I_{vd} = 1$). The decision space in time period $t \in T$ is defined as $\Phi_t = \{\mathbf{Y}_t \in \mathbf{B}^{\delta} : 0 \leq Y_{td} \leq M_{dt} - m_{d(t-1)}$ for all $d \in D\}$, where $\mathbf{B}$ denotes the binary set $\{0,1\}$. These states and decisions define the DP algorithm system dynamics: $\mathbf{S}_t = \mathbf{S}_{t-1} + \mathbf{Y}_t$. Since $\mathbf{Y}_t \in \Phi_t$ is a binary vector, a state $\mathbf{S}_t \in \Omega_t$ is accessible from state $\mathbf{S}_{t-1} \in \Omega_{t-1}$ only if $\mathbf{S}_t - \mathbf{S}_{t-1}$ is also a binary vector. Furthermore, $\mathbf{Y}_t \in \Phi_t$ being binary eliminates the necessity of the binary decision variables U_{dt}, $d \in D$, $t \in T$, because the vaccines administered in time period $t \in T$ satisfy at most one dose for a particular disease.

Given that $\mathbf{Y}_t = \mathbf{S}_t - \mathbf{S}_{t-1}$, then a transition from state $\mathbf{S}_{t-1} \in \Omega_{t-1}$ to state $\mathbf{S}_t \in \Omega_t$ requires that a dose of vaccine be administered in time period $t \in T$ for each disease in the set $D_t = \{d \in D : Y_{td} = 1\}$. The sets $V_t = \{v \in V: I_{vd} = 1$ and $d \in D_t\}$ (i.e., the set of vaccines that immunize against any disease that requires vaccination in time period $t \in T$) and D_t define a sub-instance of VFSLBP(O), where each such sub-instance is a Set-Covering problem instance, termed SCP($\mathbf{Y}_t$), with base set D_t and the collection of subsets V_t (See the Appendix for a formal definition of the Set-Covering problem). The specific Set-Covering problem instance for time period $t \in T$ and decision $\mathbf{Y}_t \in \Phi_t$ is given by

SCP($\mathbf{Y}_t$)

Minimize $\sum_{v \in V_t} c_v X_{tv}$

Subject to

$$\sum_{v \in V_t} X_{tv} I_{vd} \geq 1 \qquad \text{for all } d \in D_t,$$

$$X_{tv} \in \{0,1\} \qquad \text{for all } v \in V_t$$

To characterize the cost of decision $\mathbf{Y}_t \in \Phi_t$, which is the cost of transitioning from state $\mathbf{S}_{t-1} \in \Omega_{t-1}$ in time period $(t-1) \in T$ to state $\mathbf{S}_t \in \Omega_t$ in time period $t \in T$, define the one-period cost function $C_t(\mathbf{S}_{t-1}, \mathbf{Y}_t)$ as the cost of vaccination in time period $t \in T$ given state $\mathbf{S}_{t-1} \in \Omega_{t-1}$ and decision $\mathbf{Y}_t \in \Phi_t$. Note, however, that this one-period cost in time period $t \in T$ depends only on decision $\mathbf{Y}_t \in \Phi_t$, and hence, the optimal value of SCP($\mathbf{Y}_t$) = $C_t(\mathbf{S}_{t-1}, \mathbf{Y}_t) = C_t(\mathbf{Y}_t)$. Therefore, the optimal one-period cost over all possible decisions in time period $t \in T$ is given by $\min_{\mathbf{Y}_t \in \Phi_t} C_t(\mathbf{Y}_t)$.

Define $Z_t(\mathbf{S}_t)$ as the minimum cost of a vaccine formulary that immunizes against all diseases through time period $t \in T$ subject to the number of required doses at the end of time period $t \in T$ being equal to $\mathbf{S}_t \in \Omega_t$. Therefore, the DP optimality equation is given by the recurrence relation

$$Z_t(\mathbf{S}_t) = \min_{\mathbf{Y}_t \in \Phi_t, \mathbf{S}_{t-1} \in \Omega_{t-1} : \mathbf{S}_t = \mathbf{S}_{t-1} + \mathbf{Y}_t} \{C_t(\mathbf{Y}_t) + Z_{t-1}(\mathbf{S}_{t-1})\}.$$

Furthermore, the optimal cost of the vaccine formulary that satisfies a given childhood immunization schedule is given by

$$z^* = \min_{\mathbf{S}_\tau \in \Omega_\tau} Z_t(\mathbf{S}_t),$$

where Ω_τ is the state space for the final time period $\tau \in T$. The DP algorithm for VFSLBP(O) is now formally given.

Dynamic Programming Algorithm for VFSLBP(O)

Step 1. Initialize:

a. Initial state, $\mathbf{S}_0 \leftarrow \mathbf{0}$ (the δ-dimensional zero vector)
b. Initial cost contribution, $Z_0(\mathbf{S}_0) \leftarrow 0$
c. Set $m_{d0}, M_{d0} \leftarrow 0$ for all $d \in D$
d. Initial stage, $t \leftarrow 1$

Step 2. Compute $Z_t(\mathbf{S}_t) = \min_{\mathbf{Y}_t \in \Phi_t, \mathbf{S}_{t-1} \in \Omega_{t-1} : \mathbf{S}_t = \mathbf{S}_{t-1} + \mathbf{Y}_t} \{C_t(\mathbf{Y}_t) + Z_{t-1}(\mathbf{S}_{t-1})\}$ for each state $\mathbf{S}_t \in \Omega_t$.

Step 3. If $t < \tau$, then $t \leftarrow t + 1$ and return to Step 2. Else, stop and return $z^* = \min_{\mathbf{S}_\tau \in \Omega_\tau} Z_t(\mathbf{S}_t)$.

Example 2

This extends Example 1 for the childhood immunization schedule depicted in Figure 2. Recall the minimum dose vectors for each disease $d \in D$ are m_1 = (0,0,1,2,2,2,2,3), m_2 = (0,0,1,2,2,2,2,3), and m_3 = (0,0,0,0,0,0,0,1), where m_{dt} is the t^{th} component, $t = 1,2,\ldots,8$, of vector m_d for disease d = 1,2,3. Likewise, the maximum dose vectors for each disease $d \in D$ are M_1 = (1,2,2,2,3,3,3,3), M_2 = (0,0,1,2,3,3,3,3), and M_3 = (0,0,0,0,0,1,1,1), where M_{dt} is the t^{th} component, $t = 1,2,\ldots,8$, of vector M_d for disease d = 1,2,3. These parameters yield the following state and decision spaces:

State Space for each Time Period							
Ω_1	Ω_2	Ω_3	Ω_4	Ω_5	Ω_6	Ω_7	Ω_8
{(0,0,0), (1,0,0)}	{(0,0,0), (1,0,0), (2,0,0)}	{(1,1,0), (2,1,0)}	{(2,2,0)}	{(2,2,0), (2,3,0), (3,2,0), (3,3,0)}	{(2,2,0), (2,2,1), (2,3,0), (2,3,1), (3,2,0), (3,2,1), (3,3,0), (3,3,1)}	{(2,2,0), (2,2,1), (2,3,0), (2,3,1), (3,2,0), (3,2,1), (3,3,0), (3,3,1)}	{(3,3,1)}
Decision Space for each Time Period							
Φ_1	Φ_2	Φ_3	Φ_4	Φ_5	Φ_6	Φ_7	Φ_8
{(0,0,0), (1,0,0)}	{(0,0,0), (1,0,0)}	{(0,1,0), (1,1,0)}	{(0,1,0), (1,1,0)}	{(0,0,0), (0,1,0), (1,0,0), (1,1,0)}	{(0,0,0), (0,0,1), (0,1,0), (0,1,1), (1,0,0), (1,0,1), (1,1,0), (1,1,1)}	{(0,0,0), (0,0,1), (0,1,0), (0,1,1), (1,0,0), (1,0,1), (1,1,0), (1,1,1)}	{(0,0,0), (0,0,1), (0,1,0), (0,1,1), (1,0,0), (1,0,1), (1,1,0), (1,1,1)}.

Applying the DP algorithm, where $\Omega_0 = \{\mathbf{S}_0\} = \{(0,0,0)\}$ and $Z_0(\mathbf{S}_0) = 0$, implies for time period (stage):

t = 1:

$Z_1((0,0,0) = C_1((0,0,0)) + Z_0((0,0,0)) = 0 + 0 = 0$

$Z_1((1,0,0)) = C_1((1,0,0)) + Z_0((0,0,0)) = 1 + 0 = 1.$

Note that the value of $C_1((1,0,0))$ = SCP((1,0,0)), which is the optimal value of the IP:

Minimize $X_{11} + 3X_{13}$

Subject to

$$X_{11} + X_{13} \geq 1$$

$$X_{11}, X_{13} \in \{0,1\}.$$

The optimal value is 1 by letting $X_{11} = 1$ and $X_{13} = 0$. Continuing to the next stage yields:

$t = 2$:

$Z_2((0,0,0)) = C_2((0,0,0)) + Z_1((0,0,0)) = 0 + 0 = 0$

$Z_2((1,0,0)) = \min\{C_2((0,0,0)) + Z_1((1,0,0)); C_2((1,0,0)) + Z_1((0,0,0))\}$

$= \min\{0 + 1, 1 + 0\} = 1$

$Z_2((2,0,0)) = C_2((1,0,0)) + Z_1((1,0,0) = 1 + 1 = 2,$

$t = 3$:

$Z_3((1,1,0)) = \min\{C_3((0,1,0) + Z_2((1,0,0)); C_3((1,1,0) + Z_2((0,0,0))\}$

$= \min\{2 + 1, 3 + 0\} = 3$

$Z_3((2,1,0)) = \min\{C_3(0,1,0) + Z_2((2,0,0)); C_3((1,1,0)) + Z_2((1,0,0))\}$

$= \min\{2 + 2, 3 + 1\} = 4,$

$t = 4$:

$Z_4((2,2,0)) = \min\{C_4((1,1,0)) + Z_3((1,1,0)); C_4((0,1,0)) + Z_3((2,1,0))\}$

$= \min\{3 + 3, 2 + 4\} = 6,$

$t = 5$:

$Z_5((2,2,0)) = C_5((0,0,0)) + Z_4((2,2,0)) = 0 + 6 = 6$

$Z_5((2,3,0)) = C_5((0,1,0)) + Z_4((2,2,0)) = 2 + 6 = 8$

$Z_5((3,2,0)) = C_5((1,0,0)) + Z_4((2,2,0)) = 1 + 6 = 7$

$Z_5((3,3,0)) = C_5((1,1,0)) + Z_4((2,2,0)) = 3 + 6 = 9,$

$t = 6$:

$Z_6((2,2,0)) = C_6((0,0,0)) + Z_5((2,2,0)) = 0 + 6 = 6$

$Z_6((2,2,1)) = C_6((0,0,1)) + Z_5((2,2,0)) = 2 + 6 = 8$

$Z_6((2,3,0)) = \min\{C_6((0,0,0)) + Z_5((2,3,0)); C_6((0,1,0)) + Z_5((2,2,0))\}$

$= \min\{0 + 8, 2 + 6\} = 8$

$Z_6((3,2,0)) = \min\{C_6((0,0,0)) + Z_5((3,2,0)); C_6((1,0,0)) + Z_5((2,2,0))\}$

$= \min\{0 + 7, 1 + 6\} = 7$

$Z_6((3,3,0)) = \min\{C_6((0,0,0)) + Z_5((3,3,0)); C_6((1,0,0)) + Z_5((2,3,0)); C_6((1,1,0)) + Z_5((2,2,0)); C_6((0,1,0) + Z_5((3,2,0))\} = \min\{0 + 9, 1 + 8, 3 + 6, 2 + 7\} = 9$

$Z_6((2,3,1)) = \min\{C_6((0,1,1)) + Z_5((2,2,0)); C_6((0,0,1)) + Z_5((2,3,0))\}$

$= \min\{3 + 6, 2 + 8\} = 9$

$Z_6((3,2,1)) = \min\{C_6((1,0,1)) + Z_5((2,2,0)); C_6((0,0,1)) + Z_5((3,2,0))\}$

$= \min\{3 + 6, 2 + 7\} = 9$

$Z_6((3,3,1)) = \min\{C_6((0,0,1) + Z_5((3,3,0)); C_6((1,0,1) + Z_5((2,3,0)); C_6((1,1,1) + Z_5((2,2,0)); C_6((0,1,1) + Z_5((3,2,0))\} = \min\{2 + 9, 3 + 8, 3 + 6, 3 + 7\} = 9,$

$t = 7$:

$Z_7((2,2,0)) = C_7((0,0,0)) + Z_6((2,2,0)) = 0 + 6 = 6$

$Z_7((2,2,1)) = \min\{C_7((0,0,1)) + Z_6((2,2,0)); C_7((0,0,0)) + Z_6((2,2,1))\}$

$= \min\{2 + 6, 0 + 8\} = 8$

$Z_7((2,3,0)) = \min\{C_7((0,0,0)) + Z_6((2,3,0)); C_7((0,1,0)) + Z_6((2,2,0))\}$

$= \min\{0 + 8, 2 + 6\} = 8$

$Z_7((3,2,0)) = \min\{C_7((0,0,0)) + Z_6((3,2,0)); C_7((1,0,0)) + Z_6((2,2,0))\}$

$= \min\{0 + 7, 1 + 6\} = 7$

$Z_7((3,3,0)) = \min\{C_7((0,0,0)) + Z_6((3,3,0)); C_7((1,0,0)) + Z_6((2,3,0)); C_7((1,1,0)) + Z_6((2,2,0)); C_7((0,1,0) + Z_6((3,2,0))\} = \min\{0 + 9, 1 + 8, 3 + 6, 2 + 7\} = 9$

$Z_7((2,3,1)) = \min\{C_7((0,1,1)) + Z_6((2,2,0)); C_7((0,0,1)) + Z_6((2,3,0)); C_7((0,1,0)) + Z_6((2,2,1)); C_7((0,0,0)) + Z_6((2,3,1))\} = \min\{3 + 6, 2 + 8, 2 + 8, 0 + 9\} = 9$

$Z_7((3,2,1)) = \min\{C_7((1,0,1)) + Z_6((2,2,0)); C_7((0,0,1)) + Z_6((3,2,0)); C_7((1,0,0)) + Z_6((2,2,1)); C_7((0,0,0)) + Z_6((3,2,1))\} = \min\{3 + 6, 2 + 7, 1 + 8, 0 + 9\} = 9$

$Z_7((3,3,1)) = \min\{C_7((0,0,1) + Z_6((3,3,0)); C_7((1,0,1) + Z_6((2,3,0)); C_7((1,1,1) + Z_6((2,2,0)); C_7((0,1,1) + Z_6((3,2,0)); C_7((1,1,0) + Z_6((2,2,1)); C_7((1,0,0) + Z_6((2,3,1)); C_7((0,1,0) + Z_6((3,2,1)); C_7((0,0,0) + Z_6((3,3,1))\}$

$= \min\{2 + 9, 3 + 8, 3 + 6, 3 + 7, 3 + 8, 1 + 9, 2 + 9, 0 + 9\} = 9,$

$t = 8$:

$Z_8((3,3,1)) = \min\{C_8((0,0,1) + Z_7((3,3,0)); C_8((1,0,1) + Z_7((2,3,0)); C_7((1,1,1) + Z_7((2,2,0)); C_8((0,1,1) + Z_7((3,2,0)); C_8((1,1,0) + Z_7((2,2,1)); C_8((1,0,0) + Z_7((2,3,1)); C_8((0,1,0) + Z_7((3,2,1)); C_8((0,0,0) + Z_7((3,3,1))\}$

$= \min\{2 + 9, 3 + 8, 3 + 6, 3 + 7, 3 + 8, 1 + 9, 2 + 9, 0 + 9\} = 9.$

Therefore, the minimum cost of satisfying the childhood immunization schedule in Figure 2 is 9. Furthermore, one feasible vaccination schedule (highlighted above) that yields this optimal cost is to administer vaccine $v = 1$ in time period $t = 1$ at a cost of 1, administer vaccine $v = 2$ in time period $t = 3$ at a cost of 2, and then in time period $t = 4$, either administer the trivalent vaccine $v = 4$ or vaccines $v = 1$ and 2—either option is at a cost of 3. Finally, administer the trivalent vaccine $v = 4$ in time period $t = 6$ at a cost of 3. □

To determine the complexity of this DP algorithm, suppose that the Set-Covering problem instance with δ diseases and υ vaccines can be solved in $O(\mathbf{T}_{\text{SCP}})$ time. Furthermore, define $\mathbf{S}_{Max}$ to be the maximum number of possible states within any time period $t \in T$. Each time period requires $O((\mathbf{S}_{Max})^2 \cdot \mathbf{T}_{\text{SCP}})$ time, and hence, with τ time periods, the DP algorithm for VFSLBP(O) executes in $O(\tau(\mathbf{S}_{Max})^2 \cdot \mathbf{T}_{\text{SCP}})$ time. Since the decision problem for Set-Covering is *NP*-complete in the strong sense (Garey and Johnson 1979), a polynomial (or even pseudo-polynomial) algorithm is unlikely to exist, unless $P = NP$. The DP algorithm's worst case complexity may be improved, however, since each Set-Covering problem instance SCP($\mathbf{Y}_t$) depends only on the decision vector $\mathbf{Y}_t \in \Phi_t$. Therefore, the Set-Covering problem instance for decision $\mathbf{Y}_t \in \Phi_t$ only needs to be solved once. It can be shown that the complexity of solving for all possible decisions is $O(\upsilon\delta 2^{\delta})$. This means that for each time period $t \in T$, the complexity of Step 2 becomes $O(\delta(\mathbf{S}_{Max})^2)$, and hence, the DP algorithm has a $O(\tau\delta(\mathbf{S}_{Max})^2 + \upsilon\delta 2^{\delta})$ worst case time complexity, which is a significant improvement when $\mathbf{S}_{Max}$ is large. To exploit this added efficiency, the implementation of the DP algorithm used for the computational analysis reported in Section 3.4 employs a 'branch and remember' recursive algorithm to find the optimal cost for each Set-Covering problem instance SCP($\mathbf{Y}_t$). Therefore, SCP($\mathbf{Y}_t$) need only be computed once using the recursive algorithm *Set-Cover*. Initially, the given set of diseases for $\mathbf{Y}_t$ is D_t, and hence, $D' = D_t$.

Set-Cover(D')
    If $D' = \emptyset$ then return 0
    If Set-Covering problem for D' has been solved previously then return its optimal value
    Select a disease $d \in D'$ that requires immunization
    Let $V' = \{v \in V: I_{vd} = 1\}$ (the set of vaccines $v \in V$ that immunize against disease $d \in D'$)
    Set $BestCost = +\infty$
    For each vaccine $v \in V'$
        Let $D^* = D' \setminus \{d \in D': I_{vd} = 1\}$
        $cost = Set\text{-}Cover(D^*)$ (find the optimal cost to cover the set of diseases D^*)
        If $cost + c_v < BestCost$
            $BestCost = cost + c_v$
      Store *BestCost* for D' (remember the optimal solution for the set of diseases D')
    Return *BestCost*

Despite its exponential worst case complexity run time, the DP algorithm offers several advantages, both theoretically and computationally. First, this algorithm may be efficient in practice with the 2006 Recommended Childhood Immunization Schedule, since this schedule yields a reasonable state/decision space, and the SCP($\mathbf{Y}_t$) instances in each time period $t \in T$ are small (and, in many cases, are polynomial time solvable). Second, the DP algorithm offers insight into the theoretical structure of VFSLBP. For example, the fact that the DP algorithm yields Set-Covering problem instances in each time period allows one to exploit the theory and algorithms for Set-Covering. Third, the structure of the DP algorithm is ideal for when a child has already been partially immunized and then reenters the healthcare system to complete the immunization schedule (this problem is termed the *schedule completion problem*). Fourth, the structure of the DP algorithm makes it easier to impose restrictions that are schedule-specific by imposing such restrictions on each SCP($\mathbf{Y}_t$) instance. (See Sewell et al. (2005) for some of the restrictions that are specific to the Recommended Childhood Immunization Schedule.) Lastly, the structure of the DP algorithm is well suited for solving VFSLBP(O) related problems that include some stochastic variation. For example, during a given time period $t \in T$, a parent/guardian may refuse a particular dose of vaccine if the number of injections required is unreasonably high. Therefore, as each vaccine is administered, the probability that a parent/guardian refuses another injection increases (this problem is termed the *balking problem*).

3.3.2 Rounding Heuristics

The worst case complexity for the DP algorithm motivates the need for heuristics, which are computationally efficient and provide "good" solutions. This section presents the *Rounding* and *MAX Rounding* heuristics for VFSLBP(O)-MED. VFSLBP(O)-MED is first considered due to its simpler structure and its relation to the 2006 Recommended Childhood Immunization Schedule (all diseases have mutually exclusive doses). Both *Rounding* and *MAX Rounding* are shown to be approximation algorithms, which, by definition, execute in polynomial time and provide an approximation bound on the cost of the heuristic solution (Hochbaum 1997).

The *Rounding* and *MAX Rounding* heuristics use the solution from a linear program (LP) to construct a feasible binary solution. This technique has been applied to several other well known discrete optimization problems (Hochbaum 1997). Relaxing the binary constraint (4) for VFSLBP(O)-MED yields the LP relaxation

Minimize $\sum_{t\in T}\sum_{v\in V} c_v X_{tv}$

Subject to

$\sum_{t\in T}\sum_{v\in V} P_{djt} X_{tv} I_{vd} \geq 1$ for all $d \in D$, $j = 1,2,\ldots,n_d$

$X_{tv} \geq 0$ for all $t \in T$, $v \in V$.

Denote the optimal objective function values of VFSLBP(O)-MED and its LP relaxation as z_{IP} and z_{LP}, respectively, where $z_{LP} \leq z_{IP}$ (since the feasible region of VFSLBP(O)-MED is contained in the feasible region of its LP relaxation). Let X^*_{LP} denote the optimal decision vector for the LP relaxation, and define $\alpha_d = (\Sigma_{v\in V} I_{vd})(\max_{j=1,2,\ldots,n_d} \Sigma_{t\in T} P_{djt})$ for all diseases $d \in D$, which is the maximum number of non-zero columns in any row of the constraint matrix for VFSLBP(O)-MED corresponding to disease $d \in D$. Furthermore, define $\alpha = \max_{d\in D} \alpha_d$. The *Rounding* heuristic rounds each fractional variable in the decision vector X^*_{LP} that is greater than the threshold value $1/\alpha$. The *Rounding* heuristic is now formally given.

Rounding Heuristic for VFSLBP(O)-MED

Step 1. Solve the LP relaxation of VFSLBP(O)-MED
Step 2. $X_{tv} \leftarrow 0$ for all $t \in T$ and $v \in V$
Step 3. For all $t \in T$ and $v \in V$
    a. If $X^*_{LP_{tv}} \geq 1/\alpha$, then $X_{tv} \leftarrow 1$
Step 4. Compute and return $\Sigma_{t\in T}\Sigma_{v\in V}\ c_v X_{tv}$

Example 3

Consider the childhood immunization schedule displayed in Figure 3 together with the vaccine set $V = \{1 = \{1\}, 2 = \{2\}, 3 = \{2,3\}\}$ and cost vector $c = (2,4,4)$.

DISEASE	**TIME PERIOD**			
	1	**2**	**3**	**4**
1	Dose 1		Dose 2	
2	Dose 1			Dose 2
3		Dose 1		

Figure 3: Childhood Immunization Schedule with Mutually Exclusive Doses

In this example, z_{LP} = 16, and a feasible decision vector for $X^*_{LP} = (X^*_{LP_{11}}, X^*_{LP_{12}}, X^*_{LP_{13}}, X^*_{LP_{22}}, X^*_{LP_{23}}, X^*_{LP_{31}}, X^*_{LP_{32}}, X^*_{LP_{33}}, X^*_{LP_{41}}, X^*_{LP_{42}}, X^*_{LP_{43}})$ that yields this optimal value is X^*_{LP} = $(1, \frac{5}{6}, \frac{1}{6}, 0, 0, \frac{4}{5}, \frac{3}{4}, 0, \frac{1}{5}, \frac{1}{4}, \frac{2}{3}, \frac{1}{3})$. Furthermore, α = 4, and hence, the *Rounding* heuristic rounds all binary variables $\geq$ ¼ yielding the binary assignment (1,1,0,0,0,1,1,0,0,1,1,1), which returns an objective function value of 22. □

Lemma 3 establishes the feasibility of the solution returned by the *Rounding* heuristic.

LEMMA 3: *The Rounding heuristic for VFSLBP(O)-MED returns a feasible binary solution X,* (*i.e., a decision vector that satisfies the childhood immunization schedule*).

PROOF: Suppose the *Rounding* heuristic does not produce a feasible solution that satisfies the childhood immunization schedule. Then there exists some disease $d \in D$ whose j^{th} dose is not administered during some time period $t \in T$ such that $P_{djt} = 1$. This implies that $X^*_{LP_{tv}} < 1/\alpha$ for all decision variables in the constraint corresponding to disease $d \in D$, dose j. However, by definition of α, there are at most α decision variables in this constraint. Therefore, for disease $d \in D$, dose j,

$$\sum_{t\in T}\sum_{v\in V} P_{djt} X^*_{LP_{tv}} I_{vd} < \alpha(1/\alpha) < 1,$$

which violates the LP relaxation constraint for disease $d \in D$, dose j, but this is a contradiction, since X^*_{LP} could not be feasible. ■

Given that LP is solvable in polynomial time, it then follows that the *Rounding* heuristic executes in polynomial time. Theorem 4 shows that the cost of the binary solution returned by the *Rounding* heuristic is guaranteed to be no worse than $\alpha \cdot z_{IP}$.

THEOREM 4: *The Rounding heuristic is an* α*-approximation algorithm for VFSLBP*(*O*)–*MED*.

PROOF: Clearly, the *Rounding* heuristic executes in polynomial time since LP executes in polynomial time. It remains to show that $\Sigma_{t\in T}\Sigma_{v\in V}\ c_v X_{tv} \leq \alpha \cdot z_{IP}$. Observe that,

$$\begin{aligned}
\sum_{t\in T}\sum_{v\in V} c_v X_{tv} &\leq \sum_{t\in T}\sum_{v\in V} c_v X^*_{LP_{tv}}\,\alpha && \text{(since } X_{tv} = 1 \text{ only if } X^*_{LP_{tv}}\,\alpha \geq 1\text{)}\\
&= \alpha\, \Sigma_{t\in T}\Sigma_{v\in V}\ c_v\ X^*_{LP_{tv}} \\
&= \alpha \cdot z_{LP} \\
&\leq \alpha \cdot z_{IP} && \text{(since } z_{LP} \leq z_{IP}\text{).} \ ■
\end{aligned}$$

Theorem 4 implies some immediate corollaries for some special cases of VFSLBP(O)-MED. Corollary 1 considers a tight childhood immunization schedule such that there are at most two vaccines that immunize against each disease $d \in D$ (i.e., $\Sigma_{v\in V} I_{vd} \leq 2$ for all diseases $d \in D$), and Corollary 2 gives an upper bound on α for a tight childhood immunization schedule and for an arbitrary childhood immunization schedule.

COROLLARY 1: *Given a tight childhood immunization schedule, if $\Sigma_{v\in V} I_{vd} \leq 2$ for all diseases $d \in D$, then the Rounding heuristic is a 2-approximation algorithm for VFSLBP(O)-MED.*

PROOF: A tight childhood immunization schedule implies $\max_{j=1,2,\ldots,n_d} \Sigma_{t\in T} P_{djt} = 1$ for all diseases $d \in D$. Moreover, $\Sigma_{v\in V} I_{vd} \leq 2$ for all diseases $d \in D$, and hence, $\alpha_d = (\Sigma_{v\in V} I_{vd})(\max_{j=1,2,\ldots,n_d} \Sigma_{t\in T} P_{djt}) \leq 2$ for all diseases $d \in D$. Therefore, $\alpha = \max_{d\in D} \alpha_d \leq 2$. ■

COROLLARY 2: *Given a tight childhood immunization schedule, $\alpha \leq \upsilon$ for the Rounding heuristic, and given an arbitrary childhood immunization schedule, $\alpha \leq \upsilon\cdot\tau$ for the Rounding heuristic.*

PROOF: Given a tight childhood immunization schedule, $\max_{j=1,2,\ldots,n_d} \Sigma_{t\in T} P_{djt} = 1$ for all diseases $d \in D$, which implies $\alpha = \max_{d\in D} (\Sigma_{v\in V} I_{vd})(\max_{j=1,2,\ldots,n_d} \Sigma_{t\in T} P_{djt}) \leq \upsilon$. Moreover, for an arbitrary childhood immunization schedule, $\alpha = \max_{d\in D} (\Sigma_{v\in V} I_{vd})(\max_{j=1,2,\ldots,n_d} \Sigma_{t\in T} P_{djt}) \leq \upsilon\cdot\tau$. ■

If X^*_{LP} contains several fractional variables, then the *Rounding* heuristic tends to round too many variables to one, thereby yielding a very costly solution. Instead of rounding all variables greater than or equal to the $1/\alpha$ threshold, it seems reasonable to round only a few variables with large fractional values, since these variables are more likely to be one in the binary optimal solution. The *MAX Rounding* heuristic limits the number of rounded variables by selecting the variables with large fractional values.

To present the *MAX Rounding* heuristic, some additional notation is required. Define $\mathbf{D} = \{(d, j)\colon d \in D, j = 1,2,\ldots,n_d\}$ to be the set of all diseases ordered by dose, where $|\mathbf{D}| = \Sigma_{d=1}^{\delta} n_d$. For all time periods $t \in T$ and vaccines $v \in V$, define $C_{tv} = \{(d, j) \in \mathbf{D}\colon I_{vd} = 1$ and $P_{djt} = 1\}$, which specifies the diseases and dose that vaccine $v \in V$ immunizes against in time period $t \in T$. Therefore, $C_{tv} \subseteq \mathbf{D}$ for all time periods $t \in T$ and vaccines $v \in V$. Furthermore, in the case when all diseases $d \in D$ have mutually exclusive doses, at most one $(d, j) \in \mathbf{D}$ for all diseases $d \in D$ is contained in any set C_{tv} since, for a given disease $d \in D$ and time period

$t \in T$, P_{djt} = 1 for at most one dose j = 1,2,…,n_d, and hence, each set C_{tv} does not contain multiple doses for any disease $d \in D$. Lastly, define $f_{tv} = X^*_{LP_{tv}}$ for all time periods $t \in T$ and vaccines $v \in V$, which specifies the value of vaccine $v \in V$ in time period $t \in T$. Therefore, the *MAX Rounding* heuristic limits the number of rounded variables by greedily selecting (at each iteration) the most valuable available vaccine $v \in V$ that immunizes against the most disease doses (not yet covered) in time period $t \in T$ (i.e., rounds the variable $X^*_{LP_{tv}}$ that maximizes $f_{tv} \cdot |C_{tv}|$) until every disease dose $(d, j) \in \mathbf{D}$ is covered by some vaccine $v \in V$ in time period $t \in T$. The *MAX Rounding* heuristic is now formally given.

MAX Rounding Heuristic for VFSLBP(O)-MED

Step 1. Initialize:

a. Solve the LP relaxation of VFSLBP(O)-MED
b. $f_{tv} \leftarrow X^*_{LP_{tv}}$ for all $t \in T$, $v \in V$ such that $X^*_{LP_{tv}} \geq 1/\alpha$
c. $X_{tv} \leftarrow 0$ for all $t \in T$ and $v \in V$
d. $\hat{C}_{tv} \leftarrow C_{tv}$ for all $t \in T$ and $v \in V$

Step 2. While $\mathbf{C} = \bigcup_{\{tv: X_{tv}=1\}} C_{tv} \neq \mathbf{D}$ do

a. $(t', v') \leftarrow \arg\max_{t \in T, v \in V} f_{tv} \cdot |\hat{C}_{tv}|$ (select the non-empty set $\hat{C}_{tv}$ with the largest fractional value times the number of disease doses covered by vaccine $v \in V$ in time period $t \in T$)
b. $X_{t'v'} \leftarrow 1$ (administer vaccine $v' \in V$ in time period $t' \in T$)
c. $\hat{C}_{tv} \leftarrow \hat{C}_{tv} \backslash \hat{C}_{t'v'}$ for all $t \in T$ and $v \in V$ (remove all the disease doses covered by vaccine $v' \in V$ in time period $t' \in T$ from all remaining sets)

Step 3. Compute and return $\Sigma_{t\in T}\Sigma_{v\in V}\ c_v X_{tv}$

Example 4

Consider the childhood immunization schedule displayed in Figure 3 together with vaccine set V = {1 = {1}, 2 = {2}, 3 = {2,3}} and cost vector c = (2,4,4). Here,

$\mathbf{D}$ = {(1,1), (1,2), (2,1), (2,2), (3,1)}, and

C_{11} = {(1,1)}, C_{12} = {(2,1)}, C_{13} = {(2,1)}, C_{21} = {(1,1)}, $C_{22} = \emptyset$, C_{23} = {(3,1)},

C_{31} = {(1,2)}, $C_{32} = \emptyset$, C_{33} = {(3,1)}, C_{41} = {(1,2)}, C_{42} = {(2,2)}, C_{43} = {(2,2)}.

The *MAX Rounding* heuristic proceeds as follows:

Step 1: Initialize:

a. Solve the LP relaxation of VFSLBP(O)-MED. From the previous example, $X^*_{LP} = (1, \frac{5}{6}, \frac{1}{6}, 0, 0, \frac{4}{5}, \frac{3}{4}, 0, \frac{1}{5}, \frac{1}{4}, \frac{2}{3}, \frac{1}{3})$

b. $f_{11} = 1$, $f_{12} = \frac{5}{6}$, $f_{23} = \frac{4}{5}$, $f_{31} = \frac{3}{4}$, $f_{41} = \frac{1}{4}$, $f_{42} = \frac{2}{3}$, $f_{43} = \frac{1}{3}$

c. $X_{tv} = 0$ for $t = 1,2,3,4$ and $v = 1,2,3$

d. $\hat{C}_{11} = \{(1,1)\}$, $\hat{C}_{12} = \{(2,1)\}$, $\hat{C}_{13} = \{(2,1)\}$, $\hat{C}_{21} = \{(1,1)\}$, $\hat{C}_{23} = \{(3,1)\}$, $\hat{C}_{31} = \{(1,2)\}$, $\hat{C}_{33} = \{(3,1)\}$, $\hat{C}_{41} = \{(1,2)\}$, $\hat{C}_{42} = \{(2,2)\}$, $\hat{C}_{43} = \{(2,2)\}$

Step 2(1): **C** = $\emptyset$ since $X_{tv} = 0$ for all $t \in T$, $v \in V$, and hence

a. $(t', v') = \arg\max_{t\in T, v\in V} f_{tv} \cdot |\hat{C}_{tv}| = (1,1)$

b. $X_{11} = 1$

c. $\hat{C}_{11} = \emptyset$, $\hat{C}_{12} = \{(2,1)\}$, $\hat{C}_{13} = \{(2,1)\}$, $\hat{C}_{21} = \emptyset$, $\hat{C}_{23} = \{(3,1)\}$, $\hat{C}_{31} = \{(1,2)\}$, $\hat{C}_{33} = \{(3,1)\}$, $\hat{C}_{41} = \{(1,2)\}$, $\hat{C}_{42} = \{(2,2)\}$, $\hat{C}_{43} = \{(2,2)\}$

Step 2(2): **C** = $\{(1,1)\} \neq$ **D**

a. $(t', v') = \arg\max_{t\in T, v\in V} f_{tv} \cdot |\hat{C}_{tv}| = (1,2)$

b. $X_{12} = 1$

c. $\hat{C}_{11} = \emptyset$, $\hat{C}_{12} = \emptyset$, $\hat{C}_{13} = \emptyset$, $\hat{C}_{21} = \emptyset$, $\hat{C}_{23} = \{(3,1)\}$, $\hat{C}_{31} = \{(1,2)\}$, $\hat{C}_{33} = \{(3,1)\}$, $\hat{C}_{41} = \{(1,2)\}$, $\hat{C}_{42} = \{(2,2)\}$, $\hat{C}_{43} = \{(2,2)\}$

Step 2(3): **C** = $\{(1,1), (2,1)\} \neq$ **D**

a. $(t', v') = \arg\max_{t\in T, v\in V} f_{tv} \cdot |\hat{C}_{tv}| = (2,3)$

b. $X_{23} = 1$

c. $\hat{C}_{11} = \emptyset$, $\hat{C}_{12} = \emptyset$, $\hat{C}_{13} = \emptyset$, $\hat{C}_{21} = \emptyset$, $\hat{C}_{23} = \emptyset$, $\hat{C}_{31} = \{(1,2)\}$, $\hat{C}_{33} = \emptyset$, $\hat{C}_{41} = \{(1,2)\}$, $\hat{C}_{42} = \{(2,2)\}$, $\hat{C}_{43} = \{(2,2)\}$

Step 2(4): **C** = $\{(1,1), (2,1), (3,1)\} \neq$ **D**

a. $(t', v') = \arg\max_{t\in T, v\in V} f_{tv} \cdot |\hat{C}_{tv}| = (3,1)$

b. $X_{31} = 1$

c. $\hat{C}_{11} = \emptyset$, $\hat{C}_{12} = \emptyset$, $\hat{C}_{13} = \emptyset$, $\hat{C}_{21} = \emptyset$, $\hat{C}_{23} = \emptyset$, $\hat{C}_{31} = \emptyset$, $\hat{C}_{33} = \emptyset$, $\hat{C}_{41} = \emptyset$, $\hat{C}_{42} = \{(2,2)\}$, $\hat{C}_{43} = \{(2,2)\}$

Step 2(5): **C** = $\{(1,1), (1,2), (2,1), (3,1)\} \neq$ **D**

a. $(t', v') = \arg\max_{t \in T, v \in V} f_{tv} \cdot |\hat{C}_{tv}| = (4,2)$

b. $X_{42} = 1$

c. $\hat{C}_{11} = \emptyset,\ \hat{C}_{12} = \emptyset,\ \hat{C}_{13} = \emptyset,\ \hat{C}_{21} = \emptyset,\ \hat{C}_{23} = \emptyset,\ \hat{C}_{31} = \emptyset,\ \hat{C}_{33} = \emptyset,\ \hat{C}_{41} = \emptyset,$
$\hat{C}_{42} = \emptyset,\ \hat{C}_{43} = \emptyset$

STOP since $\mathbf{C} = \{(1,1), (1,2), (2,1), (2,2), (3,1)\} = \mathbf{D}$ and return $2X_{11} + 4X_{12} + 4X_{23} + 2X_{31} + 4X_{42} = 16$, which is the optimal cost. □

The *MAX Rounding* heuristic executes in $O(\mathbf{T}_{\text{LP}} + |\mathbf{D}|\tau\upsilon)$ time, where $\mathbf{T}_{\text{LP}}$ is the time required to solve the LP relaxation of VFSLBP(O)-MED. Furthermore, the *MAX Rounding* heuristic returns a feasible solution, since every iteration of the while loop (i.e., Step 2) administers a vaccine that satisfies at least one dose requirement for some disease $d \in D$ (i.e., every iteration covers at least one $(d, j) \in \mathbf{D}$). Moreover, the solution returned by the *MAX Rounding* heuristic can be no worse than the solution returned by the *Rounding* heuristic, and hence, the *MAX Rounding* heuristic is also an α-approximation algorithm for VFSLBP(O)-MED.

3.3.3 Primal-Dual Heuristic

This section presents the *Primal-Dual* heuristic for VFSLBP(O)-MED. In theory, a primal-dual procedure begins with a dual feasible solution and then constructs a primal solution that satisfies the Karush-Kuhn-Tucker (KKT) complementary slackness optimality conditions (Bazarra et al. 1990). If the constructed primal solution is feasible, then by the KKT conditions, it is also optimal. Otherwise a new dual feasible solution is found such that at least one additional primal variable may take on a non-zero value (again satisfying the complementary slackness conditions), and hence, results in a new primal solution. This process is repeated until the primal solution becomes feasible or the dual solution becomes unbounded, which would imply that the primal problem is infeasible (Bazaraa et al. 1990).

The *Primal-Dual* heuristic uses the sets $\mathbf{D} = \{(d, j): d \in D, j = 1,2,\ldots,n_d\}$ and $C_{tv} = \{(d, j) \in \mathbf{D}: I_{vd} = 1 \text{ and } P_{djt} = 1\}$ for all time periods $t \in T$ and vaccines $v \in V$, defined in Section 3.3.2. Observe that the set $\mathbf{D}$ together with the collection of sets C_{tv}, $t \in T$, $v \in V$, define a Set-Covering problem instance. Therefore, the following results for the *Primal-Dual* heuristic (and the *Greedy* heuristic in Section 3.3.4) are closely related to the results for the Set-Covering problem (Hochbaum 1997, Nemhauser and Wolsey 1999).

The LP relaxation of VFSLBP(O)-MED presented in Section 3.3.2 is in canonical form (i.e., min{**cx**:**Ax** $\geq$ **b**, **x** $\geq$ **0**}), and hence, the dual of the LP relaxation is of the form max{**yb**:**yA** $\leq$ **c**, **y** $\geq$ **0**} (Bazaraa et al. 1990). Let $\hat{Y}_{(d,j)}$ be the dual variable corresponding to the constraint for disease $d \in D$ and dose $j = 1,2,\ldots,n_d$ in the LP-relaxation of VFSLBP(O)-MED. The *Primal-Dual* heuristic assigns the lowest cost available vaccine for disease $d \in D$ and dose $j = 1,2,\ldots,n_d$, and then uses the vaccine cost to assign a value to the associated dual variable, $\hat{Y}_{(d,j)}$. This assignment forces at least one constraint in the dual of the LP relaxation of VFSLBP(O)-MED to become tight, which, by complementary slackness, allows the corresponding primal variable to be non-zero. The *Primal-Dual* heuristic is now formally given.

Primal-Dual Heuristic for VFSLBP(O)-MED

Step 1. Initialize:

a. $X_{tv} \leftarrow 0$ for all $t \in T$ and $v \in V$

b. $\hat{Y}_{(d,j)} \leftarrow 0$ for all $(d, j) \in \mathbf{D}$

Step 2. While $\mathbf{C} = \bigcup_{\{tv:X_{tv}=1\}} C_{tv} \neq \mathbf{D}$ do

a. Select (at random) any $(d, j) \in \mathbf{D}$ such that $(d, j) \notin \mathbf{C}$; denote this selection by (d', j')

b. $(t',v') \leftarrow \underset{t\in T, v\in V:(d',j')\in C_{tv}}{\arg\min} \left\{ c_v - \sum_{(d,j)\in C_{tv}} \hat{Y}_{(d,j)} \right\}$ (select set C_{tv} with the smallest gap between the vaccine cost and the sum of all dual variables corresponding to the disease doses covered by vaccine $v \in V$ in time period $t \in T$)

c. $\Delta_{t'v'} \leftarrow c_{v'} - \sum_{(d,j)\in C_{t'v'}} \hat{Y}_{(d,j)}$

d. $\hat{Y}_{(d',j')} \leftarrow \hat{Y}_{(d',j')} + \Delta_{t'v'}$ (estimate dual variable corresponding to $(d', j') \in \mathbf{D}$)

e. $X_{t'v'} \leftarrow 1$ (administer vaccine $v' \in V$ in time period $t' \in T$)

Step 3. Compute and return $\Sigma_{t\in T}\Sigma_{v\in V}\ c_v X_{tv}$

Example 5

Consider the childhood immunization schedule displayed in Figure 4 together with the vaccine set $V = \{1 = \{1,3\}, 2 = \{1,2,3\}\}$ and cost vector $c = (2,3)$.

DISEASE	TIME PERIOD 1	2	3	4
1	Dose 1	Dose2	Dose 3	Dose 3
2	Dose 1	Dose2	Dose 3	

3	Dose 1	Dose2	Dose 3

Figure 4: Childhood Immunization Schedule for Example 5

Here,

D = {(1,1), (1,2), (1,3), (2,1), (2,2), (2,3), (3,1), (3,2), (3,3)}, and

C_{11} = {(1,1), (3,1)}, C_{12} = {(1,1), (2,1), (3,1)}, C_{21} = {(1,2), (3,2)},

C_{22} = {(1,2), (2,2), (3,2)}, C_{31} = {(1,3), (3,3)}, C_{32} = {(1,3), (2,3), (3,3)},

C_{41} = {(1,3), (3,3)}, C_{42} = {(1,3), (3,3)}.

The *Primal-Dual* heuristic proceeds as follows:

Step 1: Initialize:

a. $X_{tv} = 0$ for $t = 1,2,3,4$ and $v = 1,2$

b. $\hat{Y}_{(d,j)} = 0$ for all $(d, j) \in \mathbf{D}$

Step 2(1): $\mathbf{C} = \emptyset$ since $X_{tv} = 0$ for all $t \in T$, $v \in V$, and hence

a. Let $(d', j') = (1,1)$

b. $(t',v') = \underset{t\in T, v\in V:(1,1)\in C_{tv}}{\arg\min} \{c_1 - 0, c_2 - 0\} = \underset{t\in T, v\in V:(1,1)\in C_{tv}}{\arg\min} \{2,3\} = (1,1)$

c. $\Delta_{11} = 2 - 0 = 2$

d. $\hat{Y}_{(1,1)} = \hat{Y}_{(1,1)} + \Delta_{11} = 0 + 2 = 2$

e. $X_{11} = 1$

Step 2(2): $\mathbf{C}$ = {(1,1), (3,1)} $\neq \mathbf{D}$

a. Let $(d', j') = (1,2)$

b. $(t',v') = \underset{t\in T, v\in V:(1,2)\in C_{tv}}{\arg\min} \{c_1 - 0, c_2 - 0\} = \underset{t\in T, v\in V:(1,2)\in C_{tv}}{\arg\min} \{2,3\} = (2,1)$

c. $\Delta_{21} = 2 - 0 = 2$

d. $\hat{Y}_{(1,2)} = \hat{Y}_{(1,2)} + \Delta_{21} = 0 + 2 = 2$

e. $X_{21} = 1$

Step 2(3): $\mathbf{C}$ = {(1,1), (1,2), (3,1), (3,2)} $\neq \mathbf{D}$

a. Let $(d', j') = (1,3)$

b. $(t',v') = \underset{t\in T, v\in V:(1,3)\in C_{tv}}{\arg\min} \{c_1 - 0, c_2 - 0\} = \underset{t\in T, v\in V:(1,3)\in C_{tv}}{\arg\min} \{2,3\} = (3,1)$

c. $\Delta_{31} = 2 - 0 = 2$

d. $\hat{Y}_{(1,3)} = \hat{Y}_{(1,3)} + \Delta_{31} = 0 + 2 = 2$

e. $X_{31} = 1$

Step 2(4): $\mathbf{C}$ = {(1,1), (1,2), (1,3), (3,1), (3,2), (3,3)} $\neq \mathbf{D}$

a. Let $(d', j') = (2,1)$

b. $(t', v') = \underset{t \in T, v \in V : (2,1) \in C_{tv}}{\arg\min} \{c_2 - 2\} = (1,2)$

c. $\Delta_{12} = 3 - 2 = 1$

d. $\hat{Y}_{(2,1)} = \hat{Y}_{(2,1)} + \Delta_{12} = 0 + 1 = 1$

e. $X_{12} = 1$

Step 2(5): $\mathbf{C} = \{(1,1), (1,2), (1,3), (2,1), (3,1), (3,2), (3,3)\} \neq \mathbf{D}$

a. Let $(d', j') = (2,2)$

b. $(t', v') = \underset{t \in T, v \in V : (2,2) \in C_{tv}}{\arg\min} \{c_2 - 2\} = (2,2)$

c. $\Delta_{22} = 3 - 2 = 1$

d. $\hat{Y}_{(2,2)} = \hat{Y}_{(2,2)} + \Delta_{22} = 0 + 1 = 1$

e. $X_{22} = 1$

Step 2(6): $\mathbf{C} = \{(1,1), (1,2), (1,3), (2,1), (2,2), (3,1), (3,2), (3,3)\} \neq \mathbf{D}$

a. Let $(d', j') = (2,3)$

b. $(t', v') = \underset{t \in T, v \in V : (2,3) \in C_{tv}}{\arg\min} \{c_2 - 2\} = (3,2)$

c. $\Delta_{32} = 3 - 2 = 1$

d. $\hat{Y}_{(2,3)} = \hat{Y}_{(2,3)} + \Delta_{32} = 0 + 1 = 1$

e. $X_{32} = 1$

STOP since $\mathbf{C} = \{(1,1), (1,2), (1,3), (2,1), (2,2), (2,3), (3,1), (3,2), (3,3)\} = \mathbf{D}$ and return $2X_{11} + 2X_{21} + 2X_{31} + 3X_{12} + 3X_{22} + 3X_{32} = 15$, which is sub-optimal. The *Primal-Dual* heuristic returns the optimal solution if the order of selection in Step 2a is $(d', j') = (2,1)$, followed by $(d', j') = (2,2)$, followed by $(d', j') = (2,3)$. □

The *Primal-Dual* heuristic executes in $O(|\mathbf{D}|\tau\upsilon)$ time, and returns a feasible solution, since every iteration of the while loop (i.e., Step 2) administers a vaccine that satisfies at least one dose requirement for some disease $d \in D$ (i.e., every iteration covers at least one $(d, j) \in \mathbf{D}$). Lemma 4 shows that the *Primal-Dual* heuristic constructs a feasible solution for the dual of the LP relaxation of VFSLBP(O)-MED.

LEMMA 4: *The Primal-Dual heuristic constructs a feasible solution $\hat{Y}$ for the dual of the LP relaxation of VFSLBP(O)-MED.*

PROOF: This is shown by induction on the while loop in Step 2 where the values for the dual variables are assigned. Initially, all the dual variables are set to zero, and hence, the base case is trivial since,

$$\sum_{(d,j)\in C_{tv}} \hat{Y}_{(d,j)} = 0 \le c_v \text{ for all } t \in T, v \in V.$$

In the induction step, assume that on iteration m of the while loop,

$$\sum_{(d,j)\in C_{tv}} \hat{Y}_{(d,j)} \le c_v \text{ for all } t \in T, v \in V.$$

Now consider iteration $m + 1$ of the while loop. During this iteration, only one dual variable is changed (namely, $\hat{Y}_{(d',j')}$). Therefore, if vaccine $v \in V$ does not immunize against disease d' $\in D$, dose j' in time period $t \in T$, then the dual constraints corresponding to vaccine $v \in V$ and time period $t \in T$ do not change (i.e., the sets C_{tv} such that $(d', j') \notin C_{tv}$ remain unchanged). However, if vaccine $v \in V$ does immunize against disease $d' \in D$, dose j' in time period $t \in T$ (i.e., $(d', j') \in C_{tv}$), then the dual variable $\hat{Y}_{(d',k')}$ is increased by $\Delta_{t'v'}$. Furthermore, all other dual variables in the constraint corresponding to vaccine $v \in V$ and time period $t \in T$ remain constant, which implies

$$\begin{aligned}
\sum_{(d,j)\in C_{tv}} \hat{Y}_{(d,j)} + \Delta_{t'v'} &= \sum_{(d,j)\in C_{tv}} \hat{Y}_{(d,j)} + \left\{ c_v - \sum_{(d,j)\in C_{t'v'}} \hat{Y}_{(d,j)} \right\} \\
&\le \sum_{(d,j)\in C_{tv}} \hat{Y}_{(d,j)} + \left\{ c_v - \sum_{(d,j)\in C_{tv}} \hat{Y}_{(d,j)} \right\} \text{ (since } (t', v') = \underset{t\in T, v\in V:(d',j')\in C_{tv}}{\arg\min} \left\{ c_v - \sum_{(d,j)\in C_{tv}} \hat{Y}_{(d,j)} \right\}) \\
&= c_v.
\end{aligned}$$

Therefore, $\sum_{(d,j)\in C_{tv}} \hat{Y}_{(d,j)} \le c_v$ for all sets C_{tv} after iteration $m + 1$. In fact, if vaccine $v \in V$ is administered in time period $t \in T$ (i.e., $X_{tv} = 1$), then $\sum_{(d,j)\in C_{tv}} \hat{Y}_{(d,j)} = c_v$, since in the algorithm when disease $d' \in D$, dose j' is covered by vaccine $v \in V$, $\Delta_{tv} = c_v - \sum_{(d,j)\in C_{tv}} \hat{Y}_{(d,j)}$ and $\hat{Y}_{(d',j')}$ is increased by exactly this amount, which causes the dual constraint to become tight. Therefore, the only primal variables permitted to take on non-zero values are those corresponding to tight dual constraints, and hence, maintains the complementary slackness condition.See the Appendix for a detailed proof. ■

Given that $\hat{Y}$ is dual feasible, then the cost of $\hat{Y}$ provides a lower bound for the cost of the optimal binary solution of VFSLBP(O)-MED, and hence, is used to bound the cost of the *Primal-Dual* heuristic solution. Therefore, the *Primal-Dual* heuristic is also an approximation algorithm for VFSLBP(O)-MED. Recall $\alpha = \max_{d \in D} \alpha_d$, where $\alpha_d = (\Sigma_{v \in V} I_{vd})(\max_{j=1,\ldots,n_d} \Sigma_{t \in T} P_{djt})$ for all diseases $d \in D$. Theorem 5 gives the approximation bound for the *Primal-Dual* heuristic.

THEOREM 5: *The Primal-Dual heuristic is an α-approximation algorithm for VFSLBP(O)-MED.*

PROOF: The *Primal-Dual* heuristic executes in $O(|\mathbf{D}|\tau\upsilon)$ time, which is clearly polynomial. Let $z_{P\text{-}D} = \Sigma_{t \in T} \Sigma_{v \in V}\ c_v X_{tv}$ be the cost of the solution returned by the *Primal-Dual* heuristic. For all $t \in T$, $v \in V$, if $X_{tv} = 1$, then, by Lemma 4, $\sum_{(d,j) \in C_{tv}} \hat{Y}_{(d,j)} = c_v$. Furthermore, define $\alpha_{(d,j)} = (\Sigma_{v \in V} I_{vd})(\Sigma_{t \in T} P_{djt})$ for all $(d, j) \in \mathbf{D}$. Therefore,

$$\begin{aligned}
z_{P\text{-}D} &= \sum_{t \in T} \sum_{v \in V} c_v X_{tv} = \sum_{t \in T} \sum_{v \in V} \sum_{(d,j) \in C_{tv}} \hat{Y}_{(d,j)} X_{tv} \\
&\leq \sum_{t \in T} \sum_{v \in V} \sum_{(d,j) \in C_{tv}} \hat{Y}_{(d,j)} \quad \text{(when } X_{tv} = 1 \text{ for all } t \in T \text{ and } v \in V) \\
&= \sum_{(d,j) \in \mathbf{D}} \hat{Y}_{(d,j)} \cdot \alpha_{(d,j)} \\
&\leq \sum_{(d,j) \in \mathbf{D}} \hat{Y}_{(d,j)} \cdot \alpha_d \quad \text{(since } \alpha_d = \max_{j=1,2,\ldots,n_d} \alpha_{(d,j)}) \\
&\leq \sum_{(d,j) \in \mathbf{D}} \hat{Y}_{(d,j)} \cdot \alpha \quad \text{(since } \alpha = \max_{d \in D} \alpha_d) \\
&\leq \alpha \cdot z_{LP} \quad \text{(by weak duality and Lemma 4)} \\
&\leq \alpha \cdot z_{IP} \quad \text{(since } z_{LP} \leq z_{IP}). \ \blacksquare
\end{aligned}$$

Corollaries 3 and 4 give similar results as those given by Corollaries 1 and 2 for the *Rounding* heuristic, except they are for the *Primal-Dual* heuristic. Corollary 3 considers a tight childhood immunization schedule such that there are at most two vaccines that immunize against each disease $d \in D$ (i.e., $\Sigma_{v \in V} I_{vd} \leq 2$ for all diseases $d \in D$), and Corrollary 4 gives an upper bound on α for a tight childhood immunization schedule and for an arbitrary childhood immunization schedule.

COROLLARY 3: *Given a tight childhood immunization schedule, if $\Sigma_{v\in V}\ I_{vd} \leq 2$ for all diseases $d \in D$, then the Primal-Dual heuristic is a 2-approximation algorithm for VFSLBP(O)-MED.*

PROOF: See proof of Corollary 1. ■

COROLLARY 4: *Given a tight childhood immunization schedule, $\alpha \leq \upsilon$ for the Primal-Dual heuristic, and given an arbitrary childhood immunization schedule, $\alpha \leq \upsilon\cdot\tau$ for the Primal-Dual heuristic.*

PROOF: See proof of Corollary 2. ■

Although the *Primal-Dual* heuristic has the same approximation bounds for VFSLBP(O)-MED as the *Rounding* and *MAX Rounding* heuristics, it should be more efficient in practice since it does not require the solution of a LP.

3.3.4 Greedy Heuristic

This section presents the *Greedy* heuristic for VFSLBP(O)-MED. The *Greedy* heuristic iteratively selects the lowest cost available vaccine that immunizes against the most disease doses. Recall, $\mathbf{D} = \{(d, j) : d \in D, j = 1,2,\ldots,n_d\}$, $C_{tv} = \{(d, j) \in \mathbf{D}: I_{vd} = 1 \text{ and } P_{djt} = 1\}$ for all time periods $t \in T$ and vaccines $v \in V$. The *Greedy* heuristic is now formally given.

Greedy Heuristic for VFSLBP(O)-MED

Step 1. Initialize:
- a. $X_{tv} \leftarrow 0$ for all $t \in T$ and $v \in V$
- b. $\hat{C}_{tv} \leftarrow C_{tv}$ for all $t \in T$ and $v \in V$

Step 2. While $\mathbf{C} = \bigcup_{\{tv: X_{tv}=1\}} C_{tv} \neq \mathbf{D}$ do
- a. $(t', v') \leftarrow \arg\min_{t\in T, v\in V} c_v / |\hat{C}_{tv}|$ (select the non-empty set $\hat{C}_{tv}$ with the smallest cost per disease doses covered by vaccine $v \in V$ in time period $t \in T$. Break ties by selecting vaccine $v \in V$ that immunizes against the most diseases in time period $t \in T$.)
- b. $X_{t'v'} \leftarrow 1$ (administer vaccine $v' \in V$ in time period $t' \in T$)
- c. $\hat{C}_{tv} \leftarrow \hat{C}_{tv} \backslash \hat{C}_{t'v'}$ for all $t \in T$ and $v \in V$ (remove all the disease doses covered by vaccine $v' \in V$ in time period $t' \in T$ from all remaining sets)

Step 3. Compute and return $\Sigma_{t\in T}\Sigma_{v\in V}\ c_v X_{tv}$

Example 6

Consider the childhood immunization schedule, vaccine set, and cost vector from Example 5. The *Greedy* heuristic proceeds as follows:

Step 1: Initialize:

a. $X_{tv} = 0$ for t = 1,2,3,4 and v = 1,2

b. $\hat{C}_{11}$ = {(1,1), (3,1)}, $\hat{C}_{12}$ = {(1,1), (2,1), (3,1)}, $\hat{C}_{21}$ = {(1,2), (3,2)}, $\hat{C}_{22}$ = {(1,2), (2,2), (3,2)}, $\hat{C}_{31}$ = {(1,3), (3,3)}, $\hat{C}_{32}$ = {(1,3), (2,3), (3,3)}, $\hat{C}_{41}$ = {(1,3),(3,3)}, $\hat{C}_{42}$ = {(1,3), (3,3)}

Step 2(1): **C** = $\emptyset$ since $X_{tv} = 0$ for all $t \in T$, $v \in V$, and hence

a. $(t', v') = \arg\min_{t \in T, v \in V} c_v / |\hat{C}_{tv}|$ = (1,2) or (2,2) or (3,2) since $|\hat{C}_{12}| > |\hat{C}_{11}|$, etc.

b. $X_{12} = 1$

c. $\hat{C}_{11} = \emptyset$, $\hat{C}_{12} = \emptyset$, $\hat{C}_{21}$ = {(1,2), (3,2)}, $\hat{C}_{22}$ = {(1,2), (2,2), (3,2)}, $\hat{C}_{31}$ = {(1,3), (3,3)}, $\hat{C}_{32}$ = {(1,3), (2,3), (3,3)}, $\hat{C}_{41}$ = {(1,3),(3,3)}, $\hat{C}_{42}$ = {(1,3), (3,3)}

Step 2(2): **C** = {(1,1), (2,1), (3,1)} ≠ **D**

a. $(t', v') = \arg\min_{t \in T, v \in V} c_v / |\hat{C}_{tv}|$ = (2,2) or (3,2) since $|\hat{C}_{22}| > |\hat{C}_{21}|$

b. $X_{22} = 1$

c. $\hat{C}_{11} = \emptyset$, $\hat{C}_{12} = \emptyset$, $\hat{C}_{21} = \emptyset$, $\hat{C}_{22} = \emptyset$, $\hat{C}_{31}$ = {(1,3), (3,3)}, $\hat{C}_{32}$ = {(1,3), (2,3), (3,3)}, $\hat{C}_{41}$ = {(1,3),(3,3)}, $\hat{C}_{42}$ = {(1,3), (3,3)}

Step 2(3): **C** = {(1,1), (1,2), (2,1), (2,2), (3,1), (3,2)} ≠ **D**

a. $(t', v') = \arg\min_{t \in T, v \in V} c_v / |\hat{C}_{tv}|$ = (3,2) since $\hat{C}_{32}$ has largest cardinality

b. $X_{32} = 1$

c. $\hat{C}_{11} = \emptyset$, $\hat{C}_{12} = \emptyset$, $\hat{C}_{21} = \emptyset$, $\hat{C}_{22} = \emptyset$, $\hat{C}_{31} = \emptyset$, $\hat{C}_{32} = \emptyset$, $\hat{C}_{41} = \emptyset$, $\hat{C}_{42} = \emptyset$

STOP since **C** = {(1,1), (1,2), (1,3), (2,1), (2,2), (2,3), (3,1), (3,2), (3,3)} = **D** and return $3X_{12} + 3X_{22} + 3X_{32} = 9$, which is the optimal cost. □

The *Greedy* heuristic executes in $O(|\mathbf{D}|\tau\upsilon)$ time, and returns a feasible solution, since every iteration of the while loop (i.e., Step 2) administers a vaccine that satisfies at least one dose requirement for some disease $d \in D$ (i.e., every iteration covers at least one $(d, j) \in \mathbf{D}$). Therefore, the *Greedy* heuristic should (in practice) be more efficient than the *MAX*

Rounding heuristic. Moreover, the *Greedy* heuristic improves the approximation bound on the returned solution. This approximation bound for the *Greedy* heuristic was derived using the same analytic approach applied to the *Primal-Dual* heuristic. To specify this approximation bound, define $\beta = \max_{t \in T, v \in V} |C_{tv}|$ and $H_k = \Sigma_{i=1}^{k} \frac{1}{i}$ (the sum of the first k elements in the harmonic series). Lemma 5 shows that the *Greedy* heuristic constructs a feasible solution to the dual of the LP relaxation of VFSLBP(O)-MED.

LEMMA 5: *The Greedy heuristic constructs a feasible solution $\hat{Y}$ for the dual of the LP relaxation of VFSLBP(O)-MED.*

PROOF: Suppose the following step is inserted in the *Greedy* heuristic between Steps 2.b and 2.c.

$$\hat{Y}_{(d,j)} \leftarrow c_{v'} /(|\hat{C}_{t'v'}| H_{\beta}) \text{ for all } (d, j) \in \hat{C}_{t'v'}.$$

This step estimates the dual variables corresponding to the disease doses covered by vaccine $v' \in V$ in time period $t' \in T$. Since the LP relaxation of VFSLBP(O)-MED is in canonical form (i.e., min{**cx**:**Ax** $\geq$ **b**, **x** $\geq$ **0**}), then the dual LP is of the form max{**yb**:**yA** $\leq$ **c**, **y** $\geq$ **0**} (Bazaraa et al. 1990). Therefore, it is necessary to show that the values assigned by the algorithm to the $\hat{Y}_{(d,j)}$ variables satisfy $\sum_{(d,j) \in C_{tv}} \hat{Y}_{(d,j)} \leq c_v$ for all $t \in T$, $v \in V$.

Consider any arbitrary set C_{tv}, $t \in T$ and $v \in V$. Let $k = |C_{tv}|$ and assume the algorithm removes elements from this set in order of the index (i.e., element 1 is the first element deleted from set C_{tv}, element 2 is the second element deleted from set C_{tv}, and so on). Without loss of generality, suppose the set **D** is reordered such that the first k elements in **D** correspond to the k elements in C_{tv}. Therefore, when the i^{th} element in **D**, denoted by (d, j), is deleted (meaning disease $d \in D$, dose j is covered by some vaccine $v' \in V$ in time period $t' \in T$), then $|\hat{C}_{tv}| \geq k - i + 1$, and hence, by the algorithm,

$$\begin{aligned}
\hat{Y}_{(d,j)} &= \frac{c_{v'}}{|\hat{C}_{t'v'}| H_{\beta}} \\
&\leq \frac{c_v}{|\hat{C}_{tv}| H_{\beta}} \qquad \left(\text{since } \frac{c_{v'}}{|\hat{C}_{t'v'}|} \leq \frac{c_v}{|\hat{C}_{tv}|} \text{ by Step 2.a of the algorithm}\right) \\
&\leq \frac{c_v}{(k-i+1)H_{\beta}} \qquad (\text{since } |\hat{C}_{tv}| \geq k - i + 1).
\end{aligned}$$

Therefore,

$$\sum_{(d,j)\in C_{tv}} \hat{Y}_{(d,j)} = \sum_{i=1}^{k} \hat{Y}_i \qquad \text{(where } i = 1,2,\ldots,k \text{ corresponds to the } i^{\text{th}} \text{ element in the set } C_{tv}\text{)}$$

$$\leq \sum_{i=1}^{k} \frac{c_v}{(k-i+1)H_\beta} = \frac{c_v}{H_\beta}\left(\frac{1}{k} + \frac{1}{k-1} + \cdots + \frac{1}{2} + 1\right) = \frac{c_v}{H_\beta} H_k$$

$$\leq c_v \qquad \text{(since } k = |C_{tv}| \leq \beta = \max_{t\in T, v\in V} |C_{tv}| \text{ implying } H_k \leq H_\beta\text{).}$$

Therefore, these values for the $\hat{Y}$ variables are a dual feasible solution to the LP relaxation of VFSLBP(O)-MED. ■

By weak duality, the cost of the dual feasible solution $\hat{Y}$ constructed by the *Greedy* heuristic provides a lower bound for the cost of the optimal binary solution of VFSLBP(O)-MED, which is used to bound the cost of the *Greedy* heuristic solution. Therefore, the *Greedy* heuristic is also an approximation algorithm for VFSLBP(O)-MED. Theorem 6 gives the approximation bound for the *Greedy* heuristic.

THEOREM 6: *The Greedy heuristic is a H_β –approximation algorithm for VFSLBP(O)-MED.*

PROOF: The *Greedy* heuristic executes in $O(|\mathbf{D}|\tau\upsilon)$ time, which is clearly polynomial. Suppose the *Greedy* heuristic requires m iterations of the while loop in Step 2, then $m \leq \tau\cdot\upsilon$. Suppose the sets C_{tv} are indexed from $1,2,\ldots,m,\ldots,\tau\cdot\upsilon$, in order of selection. Therefore, the *Greedy* heuristic selects set C_{tv_l} at iteration l. Furthermore, at iteration l, the heuristic satisfies a specific dose requirement for all diseases remaining in the set $\hat{C}_{tv_l}$ by administering vaccine $v \in V$ in time period $t \in T$. Therefore, each iteration satisfies a unique requirement for vaccine, which implies that the sets $\hat{C}_{tv_l}$, $l = 1,2,\ldots,m$ partition the set $\mathbf{D}$. Moreover, c_{v_l} = $H_\beta \sum_{(d,j)\in\hat{C}_{tv_l}} \hat{Y}_{(d,j)}$ at iteration l, since $\hat{Y}_{(d,j)} = c_{v'}/(|\hat{C}_{t'v'}|H_\beta)$ for all $(d, j) \in \hat{C}_{t'v'}$.

Therefore,

$$\sum_{t\in T}\sum_{v\in V} c_v X_{tv} = \sum_{l=1}^{m} c_{v_l} = \sum_{l=1}^{m}\left(H_\beta \sum_{(d,j)\in\hat{C}_{tv_l}} \hat{Y}_{(d,j)}\right)$$

$$= H_\beta \sum_{l=1}^{m} \sum_{(d,j)\in\hat{C}_{tv_l}} \hat{Y}_{(d,j)}$$

$$= H_\beta \sum_{(d,j)\in\mathbf{D}} \hat{Y}_{(d,j)} \qquad \text{(since the sets } \hat{C}_{tv_l} \text{ partition } \mathbf{D}\text{)}$$

$$\leq H_\beta\, z_{LP}$$

$$\leq H_\beta \, z_{IP} \qquad\qquad \text{(since } z_{LP} \leq z_{IP}\text{)}.$$

The first inequality follows from Lemma 5, where $\sum_{(d,j)\in\mathbf{D}} \hat{Y}_{(d,j)}$ is a dual feasible solution to the LP relaxation of VFSLBP(O)-MED, and by weak duality, $\sum_{(d,j)\in\mathbf{D}} \hat{Y}_{(d,j)} \leq z_{LP}$. ■

Observe that for an arbitrary childhood immunization schedule, H_β, where $\beta = \max_{t\in T, v\in V} |C_{tv}| \leq \max_{v\in V} Val(v)$, will likely be much smaller than $\alpha = \max_{d\in D} \alpha_d$. Corollary 5 considers the approximation bound for the *Greedy* heuristic using only monovalent vaccines, and Corollary 6 considers the approximation bound for the *Greedy* heuristic using bivalent and trivalent vaccines.

COROLLARY 5: *If Val*(v) = 1 *for all vaccines* $v \in V$*, then the Greedy heuristic yields the optimal vaccine formulary for VFSLBP(O)-MED.*

PROOF: If $Val(v) = 1$ for all vaccines $v \in V$, then $\beta = \max_{t\in T, v\in V} |C_{tv}| = 1$, and hence, $H_\beta = 1$. ■

COROLLARY 6: *If Val*(v) $\leq$ 2 *for all vaccines* $v \in V$*, then the Greedy heuristic is a 3/2-approximation algorithm for VFSLBP(O)-MED. Furthermore, if Val*(v) $\leq$ 3 *for all vaccines* $v \in V$*, then the Greedy heuristic is a 11/6-approximation algorithm for VFSLBP(O)-MED.*

PROOF: If $Val(v) \leq 2$ for all vaccines $v \in V$, then $\beta = \max_{t\in T, v\in V} |C_{tv}| \leq 2$, which implies that $H_\beta \leq 1 + \frac{1}{2} = \frac{3}{2}$. Similarly, if $Val(v) \leq 3$ for all vaccines $v \in V$, then $\beta = \max_{t\in T, v\in V} |C_{tv}| \leq 3$, which implies that $H_\beta \leq 1 + \frac{1}{2} + \frac{1}{3} = \frac{11}{6}$. ■

<u>Example 7</u>

The H_β approximation bound for the *Greedy* heuristic is asymptotic for some VFSLBP(O)-MED instances. To see this, consider a childhood immunization schedule with $T = \{1\}$, $D = \{1,2,3\}$, and dose vector $n = (1,1,1)$. Therefore, all diseases $d \in D$ have mutually exclusive doses and $P_{djt} = 1$ for diseases $d = 1,2,3$, dose $j = 1$, and time period $t = 1$. Let the vaccine set $V = \{1 = \{1\}, 2 = \{2\}, 3 = \{3\}, 4 = \{1,2,3\}\}$ with cost vector $c = (0.333, 0.499, 0.999, 1)$. Clearly, the *Greedy* heuristic would first select vaccine $v = 1$, then vaccine $v = 2$, and then vaccine $v = 3$ and return a cost of 0.333 + 0.499 + 0.999 = 1.831. However, the optimal solution is to administer the trivalent vaccine (i.e., vaccine $v = 4$) in time period t = 1 at a cost of c_4 = 1, and hence, the solution returned by the *Greedy* heuristic is ≈ $H_\beta \cdot z_{IP} = H_3 \cdot 1 \approx 1.833$. □

3.3.5 Generalized Heuristics

This section generalizes the *MAX Rounding*, *Primal-Dual*, and *Greedy* heuristics for VFSLBP(O)-MED by converting a VFSLBP(O) instance into two distinct VFSLBP(O)-MED instances, and then applying the *MAX Rounding*, *Primal-Dual*, and *Greedy* heuristics for each VFSLBP(O)-MED instance to find a feasible solution for the VFSLBP(O) instance.

The *MAX Rounding*, *Primal-Dual*, and *Greedy* heuristics for VFSLBP(O)-MED fail for an arbitrary VFSLBP(O) instance, where some diseases $d \in D$ in the childhood immunization schedule do not have mutually exclusive doses, since the sets C_{tv}, for all time periods $t \in T$ and vaccines $v \in V$, defined in the *MAX Rounding*, *Primal-Dual*, and *Greedy* heuristics for VFSLBP(O)-MED, no longer satisfy unique dose requirements, and since, for the diseases $d \in D$ that do not have mutually exclusive doses, there are time periods $t \in T$ when more than one required dose may be administered. For example, if vaccine $v \in V$ is a monovalent vaccine such that $I_{vd} = 1$ for disease $d \in D$, and in time period $t \in T$, P_{djt} = 1(0) for j = 1,2,(3,4,...,n_d) then $C_{tv} = \{(d,1),(d,2)\}$, and hence, administering vaccine $v \in V$ in time period $t \in T$ satisfies doses 1 and 2 for disease $d \in D$. Therefore, to ensure the sets C_{tv} satisfy unique dose requirements, consider two variations of the set C_{tv} for all time periods $t \in T$ and vaccines $v \in V$

1) *Minimum Dose*: $C_{tv}^{MIN} = \{(d, k) \in \mathbf{D}: I_{vd} = 1 \text{ and } k = \min\{j: P_{djt} = 1\}\}$
2) *Maximum Dose*: $C_{tv}^{MAX} = \{(d, k) \in \mathbf{D}: I_{vd} = 1 \text{ and } k = \max\{j: P_{djt} = 1\}\}$.

Variation 1) or 2) ensure that set C_{tv} (i.e., $C_{tv} = C_{tv}^{MIN}$ for all time periods $t \in T$ and vaccines $v \in V$, or $C_{tv} = C_{tv}^{MAX}$ for all time periods $t \in T$ and vaccines $v \in V$) satisfies unique dose requirements for all diseases $d \in D$, and hence, each variation converts a VFSLBP(O) instance into a distinct VFSLBP(O)-MED instance.

Therefore, the *A* heuristic for VFSLBP(O) converts a VFSLBP(O) instance into two distinct VFSLBP(O)-MED instances, and executes the *A* heuristic for VFSLBP(O)-MED on each distinct VFSLBP(O)-MED instance, where *A* is the *MAX Rounding*, *Primal-Dual*, or *Greedy* heuristic. The *A* heuristic is now formally given.

A Heuristic for VFSLBP(O)

Step 1. Select $A \in \{MAX\ Rounding, Primal\text{-}Dual, Greedy\}$
Step 2. Initialize:
  a. Let $\mathbf{D} = \{(d, j) : d \in D, j = 1,2,\ldots,n_d\}$ and $C_{tv} = C_{tv}^{MIN}$ for all $t \in T, v \in V$
Step 3. Execute *A heuristic for VFSLBP(O)-MED* and return $Z_{MIN} = \Sigma_{t\in T}\Sigma_{v\in V}\ \ c_v X_{tv}$
Step 4. Initialize:
  a. Let $\mathbf{D} = \{(d, j) : d \in D, j = 1,2,\ldots,n_d\}$ and $C_{tv} = C_{tv}^{MAX}$ for all $t \in T, v \in V$
Step 5. Execute *A heuristic for VFSLBP(O)-MED* and return $Z_{MAX} = \Sigma_{t\in T}\Sigma_{v\in V}\ \ c_v X_{tv}$
Step 6. Return $\min\{Z_{MIN}, Z_{MAX}\}$

The *A* heuristic executes in $O(|\mathbf{D}|\tau\upsilon)$ time for *A* = *Primal-Dual* or *Greedy* and $O(|\mathbf{T}_{\mathrm{LP}} + |\mathbf{D}|\tau\upsilon)$ time for *A* = *MAX Rounding*, where $\mathbf{T}_{\mathrm{LP}}$ is the time required to solve the LP relaxations of both distinct VFSLBP(O)-MED instances.

Furthermore, the *A* heuristic returns a feasible solution for VFSLBP(O) provided a restriction is placed on the given childhood immunization schedule. To describe this restriction, dose $j = 1,2,\ldots,n_d$ is said to *dominate* dose $k = 1,2,\ldots,n_d$, $j \neq k$, for disease $d \in D$ if $P_{djt} \geq P_{dkt}$ for all time periods $t \in T$. If disease $d \in D$ has no dominant doses, then the time periods when dose $j = 1,2,\ldots,n_d$ may be administered do not completely overlap with the time periods when dose $k = 1,2,\ldots,n_d$, $j \neq k$, may be administered, and hence, for all $j = 1,2,\ldots,(n_d - 1)$, there exists time periods t, $t' \in T$ such that $P_{djt} = 1$ and $P_{d(j+1)t} = 0$ and $P_{djt'} = 0$ and $P_{d(j+1)t'} = 1$. Therefore, for VFSLBP(O), a restriction placed on the childhood immunization schedule is that, for all diseases $d \in D$, dose $j = 1,2,\ldots,n_d$ does not dominate dose $k = 1,2,\ldots,n_d$, $j \neq k$. All of the diseases in the 2006 Recommended Childhood Immunization Schedule do not have a dose that dominates any other dose, and future schedules should also meet this restriction, since there is a biological spacing requirement between each dose of vaccine for every disease $d \in D$.

This restriction ensures every $(d, j) \in \mathbf{D}$ (in Steps 2.a and 4.a) is contained in some set C_{tv} for at least one time period $t \in T$ and vaccine $v \in V$. Therefore, the *A* heuristic returns a feasible solution for VFSLBP(O) (assuming VFSLBP(O) has a feasible solution), since every iteration of the *A* heuristic for VFSLBP(O)-MED (in Step 3 and Step 5) administers a vaccine that satisfies at least one dose requirement for some disease $d \in D$ (i.e., every iteration covers at least one $(d, j) \in \mathbf{D}$).

The approximation bounds shown for the *A* heuristic for VFSLBP(O)-MED do not apply to the solution returned by the *A* heuristic for VFSLBP(O) (in Step 6), since the *A* heuristic for VFSLBP(O) converts a VFSLBP(O) instance into two distinct VFSLBP(O)-MED instances. Work is in progress to determine approximation bounds for the *A* heuristic for VFSLBP(O).

3.4 Computational Results

This section reports computational results comparing the *MAX Rounding*, *Primal-Dual*, and *Greedy* heuristics and the DP algorithm presented in Section 3.3.1. Computational results are also reported for an IP branch and bound (IP B&B) algorithm. However, the *Rounding* heuristic results are not reported, since this heuristic (empirically) yields very costly solutions. The *MAX Rounding*, *Primal-Dual*, and *Greedy* heuristics and the DP and IP B&B algorithms were executed on three sets of test problems to demonstrate their computational effectiveness and limitations. The first test problem is the 2006 Recommended Childhood Immunization Schedule. The second set of test problems are randomly generated based on hypothetical future childhood immunization schedules, while the third set of test problems are larger, randomly generated childhood immunization schedules executed with several different vaccine sets. The size of these randomly generated childhood immunization schedules assume that the future Recommended Childhood Immunization Schedules will expand to include more diseases and time periods, and hence, will require a larger number of both monovalent and combination vaccines. These assumptions are reasonable, given recent trends in expanding the schedule. For example, four time periods and three diseases have been added to the Recommended Childhood Immunization Schedule since 1995, and there are currently several vaccine products being marketed and tested for use in children (CDC 1995, Cochi 2005, Infectious Diseases in Children 2002).

In this section, the solution quality effectiveness measure θ is reported for each heuristic, where $\theta = Z_{Heuristic} / Z^*$ and $Z_{Heuristic}$ is the objective function cost returned by the heuristic and Z^* is the optimal objective function cost (returned by the exact algorithms). The execution time (in CPU seconds) is also reported for each heuristic and exact algorithm, which is the efficiency effectiveness measure. All heuristics and exact algorithms were coded and executed in *MATLAB*v7.0 on a 2.4 MHz Pentium IV with 1GB of RAM including the IP B&B algorithm (using default settings) from *MATLAB*'s optimization toolbox.

The first test problem is the 2006 Recommended Childhood Immunization Schedule displayed in Figure 1. Therefore, D = {1 = Hepatitis B, 2 = Diphtheria-Tetanus-Pertussis, 3 = *Haemophilus influenzae* type b, 4 = Polio, 5 = Measles-Mumps-Rubella, 6 = Varicella, 7 = Pneumococcus, 8 = Influenza, 9 = Hepatitis A} with dose vector $n = (3, 5, 4, 4, 2, 1, 4, 1, 2)$, since diphtheria, tetanus, and pertussis are considered one disease and measles, mumps, and rubella are also considered one disease, and $T = \{1,2,\ldots,10\}$. The vaccine set is V = {1 = {1}, 2 = {2}, 3 = {3}, 4 = {4}, 5 = {5}, 6 = {6}, 7 = {7}, 8 = {8}, 9 = {9}, 10 = {2,3}, 11 = {1,3}, 12 = {1,2,4}}. The parameters I_{vd} are indicated by the set V. For example, vaccine 1 is the monovalent vaccine for disease 1 (Hepatits B) and vaccine 12 is the combination vaccine *Pediarix*® that immunizes against diseases 1 (Hepatitis B), 2 (Diphtheria-Tetanus-Pertussis), and 4 (Polio). The schedule parameters P_{djt}, Q_{dt}, and m_{dt} for diseases $d \in D$, dose $j = 1,2,\ldots,n_d$, and time periods $t \in T$ are all obtained from Figure 1. For example, for disease $d = 1$ = Hepatitis B and dose $j = 2$, $P_{djt} = 1(0)$ for time periods $t = 2,3(1,4,5,6,7,8,9,10)$. Three different cost scenarios are evaluated. The first scenario only considers the actual purchase price of the vaccines. In particular, the cost vector c = (9.00, 12.75, 7.66, 10.42, 16.67, 52.25, 54.12, 9.71, 12.10, 24.62, 24.50, 38.34), where c_v, $v = 1,2,\ldots,12$, is the Federal contract purchase price (in US$) for vaccine $v \in V$ (CDC Vaccine Price List 2005). The second scenario includes the purchase price of the vaccine and a fixed injection cost of $10/injection, and the final scenario includes the purchase price, the fixed injection cost, and a preparation cost of $3/injection. Table 2 reports the objective function cost Z and execution time (in CPU seconds) for each heuristic and exact algorithm and for each scenario. Table 2 also reports the solution quality effectiveness measure θ for each heuristic.

Table 2: Computational Results for 2006 Recommend Childhood Immunization Schedule

	Scenario 1			**Scenario 2**			**Scenario 3**		
Algorithm	***Z***	**Time**	***θ***	***Z***	**Time**	***θ***	***Z***	**Time**	***θ***
MAX Rounding	499.05	0.13	1.00	736.77	0.13	1.02	796.77	0.13	1.02
Primal-Dual	499.05	0.06	1.00	910.65	0.03	1.27	988.65	0.08	1.27
Greedy	499.05	0.06	1.00	719.81	0.05	1.00	779.81	0.05	1.00
DP	499.05	0.32		719.81	0.30		779.81	0.31	
IP B&B	499.05	0.91		719.81	0.92		779.81	0.92	

Lemma 1 and Theorem 3 showed that VFSLBP(O) is polynomial time solvable when all vaccines $v \in V$ are monovalent vaccines, and hence, the results reported in Table 2 for Scenario 1 are not surprising, given that most vaccines $v \in V$ are monovalent. In fact, the combination vaccines are not competitively priced when considering the purchase price alone. The fixed costs considered in Scenarios 2 and 3 penalize the monovalent vaccines and make the combination vaccines more economical. For example, the purchase prices for monovalent vaccines 1, 2, and 4 sum to \$32.17, which is less than the \$38.34 purchase price for the combination vaccine *Pediarix®* (i.e., v = 12). However, in Scenario 2, the total cost of the combination vaccine *Pediarix®* is \$48.34, whereas the total costs for monovalent vaccines 1, 2, and 4 sum to \$62.17. Observe that the exact algorithms were very efficient. However, as the next set of test problems illustrate, this is unlikely to occur for future Recommended Childhood Immunization Schedules, as the schedule expands and more combination vaccines enter the market.

The second set of test problems considers hypothetical future childhood immunization schedules. Each heuristic and exact algorithm were executed on 60 randomly generated childhood immunization schedules (30 schedules with mutually exclusive doses (MED) for all diseases $d \in D$ and 30 schedules with non-mutually exclusive doses (non-MED)) with 15 time periods, 75 vaccines, and 11 diseases. Therefore, each random childhood immunization schedule reflects a gradual expansion in the sets D (from 9 to 11 diseases) and T (from 10 to 15 time periods) and a significant increase in the number of available vaccines, particularly, combination vaccines. In each random childhood immunization schedule, $1 \leq n_d \leq 5$ for all disease $d \in D$, $1 \leq Val(v) \leq 6$ and $c_v \sim U(10,80)$ (uniformly distributed) for all vaccines $v \in V$, and $P_{djt} = 1$ for at most three time periods $t \in T$ for every disease $d \in D$ and dose $j = 1,2,\ldots,n_d$. Table 3 reports the average μ and the standard deviation σ for the execution time (in CPU seconds) and solution quality θ for each type of schedule (MED and non-MED) averaged over the 30 random childhood immunization schedules.

Table 3: Computational Results for Future Childhood Immunization Schedule

	Schedule Type							
	MED				**non-MED**			
	Time		**θ**		**Time**		**θ**	
Algorithm	**μ**	**σ**	**μ**	**σ**	**μ**	**σ**	**μ**	**σ**
MAX Rounding	0.42	0.05	1.07	0.06	0.97	0.13	1.07	0.07
Primal-Dual	0.14	0.02	1.68	0.22	0.31	0.04	1.63	0.25
Greedy	0.24	0.03	1.06	0.07	0.57	0.07	1.10	0.08
DP	3.1	1.2			10.6	3.5		
IP B&B	40.6	46.5			3572(1755)[a]	2994(1808)		

[a]The IP B&B algorithm found the optimal solution for 20 of the 30 non-MED childhood immunization schedules, but exceeded the default execution time limit (two hours) for the remaining three childhood immunization schedules. The statistics shown in parentheses are for the 20 childhood immunization schedules for which the IP B&B algorithm found the optimal solution.

The solutions returned by the *MAX Rounding* and *Greedy* heuristics on average were within ten percent of the optimal solution. Across all 60 randomly generated childhood immunization schedules, the optimal solution was returned 24 times (15 MED schedules, 9 Non-MED schedules) by at least one of these two heuristics. The *Primal-Dual* heuristic was the most efficient, but returned the poorest quality solutions. Moreover, the heuristics for the non-MED random childhood immunization schedules required longer execution times than for the MED random childhood immunization schedules, since for the non-MED random childhood immunization schedules, each heuristic must be implemented twice—once for the minimum dose definition of the sets C_{tv} and once for the maximum dose definition. The exact algorithms required significantly more time to execute than the heuristics (i.e., the least efficient heuristic *MAX Rounding* was 7 times (14 times) faster than the most efficient exact algorithm DP for MED (Non-MED) schedules). Moreover, the IP B&B algorithm, on average, required significantly more execution time than the DP algorithm. Furthermore, the DP algorithm showed far less variability in its execution time. The observed difference in execution time between the heuristics and exact algorithms reported in Table 3 could be problematic for practical uses. For example, a webpage used to find a "good" vaccine formulary for a given childhood immunization schedule would require an algorithm to execute in real-time, since most web users would terminate a web application that required several seconds or minutes to execute. Moreover, the difference in execution time between the heuristics and exact algorithms will provide an efficient analysis of larger childhood immunization schedules that may involve Monte Carlo simulation (see Jacobson and Sewell 2002) or the balking problem (described in Section 3.3.1), where either of these may require the solution of hundreds of thousands of VFSLBP(O) instances. Furthermore, the childhood immunization schedule may need to be solved for each child, on a case by case basis, and hence, efficient algorithms are needed in order to provide, in real-time, practical value for the public health community.

The third set of test problems considers larger randomly generated childhood immunization schedules that demonstrate the effect of combination vaccines and further demonstrate how the schedule's size affects the efficiency and solution quality of each heuristic and exact algorithm. Each heuristic and exact algorithm were executed on 30 randomly generated childhood immunization schedules (with non-MED) with 20 time periods, 90 vaccines, and 13 diseases such that $1 \le n_d \le 5$ for all diseases $d \in D$, $c_v \sim U(10,80)$ for all vaccines $v \in V$, and $P_{djt} = 1$ for at most four time periods $t \in T$ for every disease $d \in D$ and dose $j = 1,2,\ldots,n_d$. For each randomly generated childhood immunization schedule, each heuristic and exact algorithm was executed six times, where in execution $i = 1,2,\ldots,6$, $Val(v) \le i$ for all vaccines $v \in V$. Table 4 reports for each heuristic, algorithm and valency, the execution time (in CPU seconds), and the measure θ averaged across all 30 randomly generated childhood immunization schedules. The overall average and standard deviation across all vaccine sets is also reported for each heuristic and exact algorithm as well as the relative LP-IP gap = Z_{IP}/Z_{LP}, where Z_{IP} (Z_{LP}) is the optimal objective function cost for the IP (LP relaxation).

Table 4: Computational Results for the Effect of Combination Vaccines

	Heuristics						Exact Algorithms		
	MAX Round		*Primal -Dual*		*Greedy*		DP	IP B&B[a]	LP-IP
***Val*(*v*) ≤**	**Time**	θ	**Time**	θ	**Time**	θ	**Time**	**Time**	**GAP**
1	2.8	1.00	1.0	1.00	2.4	1.00	56	5	1.00 (1.00)
2	1.8	1.02	0.8	1.42	1.6	1.06	61	164	1.01 (1.01)
3	1.6	1.02	0.7	1.59	1.3	1.07	64	861 (156)	1.01 (1.01)
4	1.5	1.03	0.6	1.64	1.1	1.08	68	2303 (521)	1.02 (1.01)
5	1.4	1.07	0.6	1.68	1.0	1.08	71	3999 (1545)	1.04 (1.02)
6	1.4	1.08	0.5	1.59	0.9	1.08	75	5493 (1488)	1.06 (1.02)
Average	1.8	1.04	0.7	1.48	1.4	1.06	66	2138 (647)	1.02 (1.01)
St Dev	0.5	0.03	0.2	0.25	0.6	0.03	7	2228 (695)	0.02 (0.01)

[a]The IP B&B algorithm found the optimal solution for 135 of the 180 VFSLBP(O) instances (each childhood immunization schedule was executed six times), but exceeded the default execution time limit (two hours) for the remaining 45 instances (3 instances for $Val(v) \le 3$, 8 instances for $Val(v) \le 4$, 13 instances for $Val(v) \le 5$, and 21 instances for $Val(v) \le 6$). The statistics shown in parentheses are for the 135 instances for which the IP B&B algorithm found the optimal solution

The data reported in Table 4 shows that in most cases the execution time required to optimally solve VFSLBP(O) steadily increased as the valency of the vaccine set increased, while the execution times for the heuristics decreased as the valency of the vaccine set increased. For example, when $Val(v) = 1$ for all vaccines $v \in V$, the least efficient heuristic *MAX Rounding* only slightly outperformed the most efficient exact algorithm IP B&B (2.8 vs. 5 seconds). However, when $Val(v) \le 5(6)$ for all vaccines $v \in V$, the least efficient heuristic *MAX Rounding* was 50 times (54 times) faster than the most efficient exact algorithm DP. As expected, however, the heuristic solution quality deteriorated as the valency of the vaccine set increased, but still fell within ten percent of the optimal solution for the *MAX Rounding* and *Greedy* heuristics. The observed decrease in execution time for the heuristics is intuitive since each heuristic should require less iterations when more vaccines are able to protect against multiple diseases. In the case of monovalent vaccines, each heuristic found the optimal solution on every random childhood immunization schedule, which is consistent with the complexity results and approximation bounds presented earlier in this chapter.

The data reported in Tables 2-4 all suggest that on average, the DP algorithm requires significantly less computational effort to find the optimal solution than is required by the IP B&B algorithm and with less variability. As shown in Section 3.3.1, the computational complexity of the DP algorithm is highly sensitive to the number of diseases, since the decision space is bounded above by 2^{δ} and $\mathbf{S}_{Max}$ also depends on the number of diseases. On the other hand, the computational complexity of the IP B&B algorithm is highly sensitive to the number of decision variables, since the number of possible branches is bounded above by $2^{\tau \cdot \upsilon}$ (there are $\tau \cdot \upsilon$ binary decision variables in an instance of VFSLBP(O)). Furthermore, the computational effort of an IP B&B algorithm is sensitive to the gap between the cost of the optimal integer solution and the corresponding cost of the optimal LP relaxation solution, since a large gap would tend to require more branching to find the optimal integer solution (Nemhauser and Wolsey 1999). The data reported in Table 2-4 provide empirical support for these remarks. For example, the DP algorithm was always more efficient than the IP B&B algorithm except in Table 4 when $Val(v) = 1$ and $\delta = 13$, and the IP B&B algorithm exceeded the default execution time limit (two hours) for 45 of the 180 VFSLBP(O) instances (each childhood immunization schedule was executed six times) when the average LP-IP gap was 1.06, which is two standard deviations above the average LP-IP gap. Therefore, it is reasonable to conjecture that the DP algorithm will be more efficient for childhood immunization schedules when the number of diseases remains relatively small (i.e., $\delta < 15$), and the IP B&B algorithm will be more efficient for childhood immunization schedules when the number of diseases is large (i.e., $\delta \geq 15$) and the number of time periods and vaccines remains reasonable (i.e., $\tau \cdot \upsilon < 1500$).

Chapter 4: The Vaccine Formulary Selection with Restricted Extraimmunization Problem

This chapter presents a general model that limits the amount of extraimmunization for any given childhood immunization schedule. It also rigorously explores the theoretical structure of this general model and provides an extensive computational study is also presented. The chapter is organized as follows. Section 4.1 presents general models (formulated as a decision problem and as a discrete optimization problem) that determine the set of vaccines (i.e., a vaccine formulary) that should be used in a clinical environment to satisfy any given childhood immunization schedule while restricting extraimmunization. Section 4.2 presents the computational complexity of the decision/discrete optimization problems. Section 4.3 presents a description and analysis of several algorithms, both exact and heuristic, for solving the discrete optimization problem. Finally, Section 4.4 presents a computational comparison of these algorithms.

4.1 Model Formulation and Terminology

A model formulation for a decision problem and a discrete optimization problem used to design a vaccine formulary that satisfies a given childhood immunization schedule while restricting extraimmunization is presented. Some simplifications and extensions of the discrete optimization problem are also described.

Given a childhood immunization schedule, the decision problem, termed the Vaccine Formulary Selection with Restricted Extraimmunization Problem (VFSREP), asks whether it is possible to design a vaccine formulary that restricts extraimmunization for a specified set of diseases. This problem is now formally stated.

Vaccine Formulary Selection with Restricted Extraimmunization Problem (VFSREP)

Given:

- A set of time periods, $T = \{1,2,\ldots,\tau\}$,
- a set of diseases, $D = \{1,2,\ldots,\delta\}$,
- a set of diseases where extraimmunization is permitted, $D_E \subseteq D$, with $|D_E| = \delta_E$,

- a set of diseases where extraimmunization is not permitted, $D_{NE} = D \backslash D_E$, with $|D_{NE}| = \delta_{NE}$,
- a set of vaccines $V = \{1,2,\ldots,\upsilon\}$, available to be administered to immunize against the δ diseases,
- the number of doses of a vaccine that must be administered for immunization against the δ diseases, $n_1,n_2,\ldots, n_\delta$,
- a set of binary parameters that indicate which vaccines immunize against which diseases; therefore, $I_{vd} = 1$ if vaccine $v \in V$ immunizes against disease $d \in D$, 0 otherwise,
- a set of binary parameters that indicate the set of time periods in which a particular dose of a vaccine may be administered to immunize against a disease; therefore, $P_{djt} = 1$ if in time period $t \in T$, a vaccine may be administered to satisfy the j^{th} dose, $j = 1,2,\ldots,n_d$, requirement for disease $d \in D$, 0 otherwise,
- a set of binary parameters that indicate the set of time periods in which a vaccine may be administered to satisfy any dose requirement against a disease; therefore, $Q_{dt} = 1$ if in time period $t \in T$, a vaccine may be administered to satisfy any dose requirement against disease $d \in D$, 0 otherwise, (i.e., for any disease $d \in D$ and time period $t \in T$, $Q_{dt} = 1$ if and only if $P_{djt} = 1$ for some dose $j = 1,2,\ldots,n_d$),
- a set of binary parameters that indicate the set of time periods in which no dose of a vaccine may be administered to immunize against a disease where extraimmunization is not permitted; therefore, $R_{dt} = 1$ if in time period $t \in T$, no dose of a vaccine may be administered to immunize against disease $d \in D_{NE}$, 0 otherwise, (i.e., for any disease $d \in D_{NE}$ and time period $t \in T$, $R_{dt} = 1$ if and only if $Q_{dt} = 0$),
- a set of integer parameters that indicate the minimum number of doses of a vaccine required for disease $d \in D$ through time period $t \in T$; denoted by m_{dt}.

Question: Does there exist a set of vaccines from V that can be administered over the time periods in T such that these vaccines immunize against all the diseases in D while restricting extraimmunization, (i.e., do there exist values for the binary decision variables X_{tv}, $t \in T$, $v \in V$, where $X_{tv} = 1$ if vaccine $v \in V$ is administered in time period $t \in T$, 0 otherwise, and for the binary variables U_{dt}, $d \in D$, $t \in T$, where $U_{dt} = 1$ if any vaccine $v \in V$ that immunizes against disease $d \in D$ is administered in time period $t \in T$

, 0 otherwise, such that for all diseases $d \in D$, $\Sigma_{t\in T}\Sigma_{v\in V} P_{djt} X_{tv} I_{vd} \geq 1$ for dose $j = 1,2,\ldots,n_d$, $\Sigma_{t=1,2,\ldots,t'} Q_{dt} U_{dt} \geq m_{dt'}$ for all time periods $t' \in T$, and $\Sigma_{v\in V} Q_{dt} X_{tv} I_{vd} \geq U_{dt}$ for all time periods $t \in T$, and for all diseases $d \in D_{NE}$, $\Sigma_{t\in T}\Sigma_{v\in V} R_{dt} X_{tv} I_{vd} = 0$ and $\Sigma_{t\in T} \Sigma_{v\in V} Q_{dt} X_{tv} I_{vd} = n_d$)?

In the VFSREP formulation, the given sets and parameters equate to a childhood immunization schedule. As was the case for VFSLBP in Chapter 3, the doses for all diseases $d \in D$ are assumed to be *sequentially ordered*, and the dose (n_d and m_{dt}) and schedule (P_{djt} and Q_{dt}) parameters are also analogous to those in VFSLPB. Note that schedule parameters P_{djt} and Q_{dt} specify the time periods when vaccination is "permitted" (or useful) for disease $d \in D$, while the schedule parameters R_{dt} specify the time periods when vaccination is "restricted" for disease $d \in D_{NE}$. For example, assuming disease d = hepatitis B $\in D_{NE}$, Figure 1 implies $Q_{dt} = 1(0)$ for time periods t =1,2,3,5,6,7,8(4,9,10) and $R_{dt} = 1(0)$ for time periods t = 4,9,10(1,2,3,5,6,7,8). The set D_{NE} is the set of diseases where extraimmunization is restricted based on biological and/or philosophical constraints, and hence, may change for each child, on a case-by-case basis. The question in VFSREP asks if there exists a vaccine formulary administered over the time periods in T that satisfies the given childhood immunization schedule and restricts extraimmunization for the diseases in the set D_{NE} (i.e., a variable assignment for the binary decision variables X_{tv}, for all time periods $t \in T$ and vaccines $v \in V$, and for the binary decision variables U_{dt}, for all diseases $d \in D$ and time periods $t \in T$, that satisfies the per dose requirements ($\Sigma_{t\in T}\Sigma_{v\in V} P_{djt} X_{tv} I_{vd} \geq 1$ for dose $j = 1,2,\ldots,n_d$) and total dosage requirements ($\Sigma_{t=1,2,\ldots,t'} Q_{dt} U_{dt} \geq m_{dt'}$ for all time period $t' \in T$ and $\Sigma_{v\in V} Q_{dt} X_{tv} I_{vd} \geq U_{dt}$ for all time periods $t \in T$) for each disease $d \in D$, and does not exceed the total dosage requirements ($\Sigma_{t\in T} \Sigma_{v\in V} Q_{dt} X_{tv} I_{vd} = n_d$) or provide a dose in a time period when no dose of a vaccine may be administered ($\Sigma_{t\in T}\Sigma_{v\in V} R_{dt} X_{tv} I_{vd} = 0$) for each disease $d \in D_{NE}$).

A solution to VFSREP may be determined by solving a discrete optimization problem. To describe this problem, some additional parameters and variables are needed. Let

- $\rho_d \in \mathbf{Q}^+$ be the weight of extraimmunization for disease $d \in D_{NE}$ for all time periods $t \in T$ such that $Q_{dt} = 1$ (i.e., in time periods when vaccination is permitted),

- $\gamma_d \in \mathbf{Q}^+$ be the weight of extraimmunization for disease $d \in D_{NE}$ for all time periods $t \in T$ such that $R_{dt} = 1$ (i.e., in time periods when vaccination is restricted),
- $Z_d^P \in \mathbf{Z}^+ \cup \{0\}$ be the number of extra doses of vaccine administered for disease $d \in D_{NE}$ in all time periods $t \in T$ such that $Q_{dt} = 1$, (i.e., in time periods when vaccination is permitted), and
- $Z_d^R \in \mathbf{Z}^+ \cup \{0\}$ be the number of extra doses of vaccine administered for disease $d \in D_{NE}$ in all time periods $t \in T$ such that $R_{dt} = 1$, (i.e., in time periods when vaccination is restricted),

where $\mathbf{Q}^+$ and $\mathbf{Z}^+ \cup \{0\}$ correspond to the set of all positive rational numbers and the set of all non-negative integers, respectively. Therefore, the following integer program can be used to answer VFSREP.

Integer Programming Model for Vaccine Formulary Selection with Restricted Extraimmunization Problem (VFSREP(O))

$$\text{Minimize} \quad \sum_{d \in D_{NE}} \rho_d Z_d^P + \gamma_d Z_d^R \tag{O}$$

Subject to

$$\sum_{t \in T} \sum_{v \in V} P_{djt} X_{tv} I_{vd} \geq 1 \qquad \text{for all } d \in D,\ j = 1,2,\ldots,n_d, \tag{1}$$

$$\sum_{t=1,2,\ldots,t'} Q_{dt} U_{dt} \geq m_{dt'} \qquad \text{for all } d \in D,\ t' \in T, \tag{2}$$

$$\sum_{v \in V} Q_{dt} X_{tv} I_{vd} \geq U_{dt} \qquad \text{for all } d \in D,\ t \in T, \tag{3}$$

$$\sum_{t \in T} \sum_{v \in V} Q_{dt} X_{tv} I_{vd} - Z_d^P = n_d \qquad \text{for all } d \in D_{NE}, \tag{4}$$

$$\sum_{t \in T} \sum_{v \in V} R_{dt} X_{tv} I_{vd} - Z_d^R = 0 \qquad \text{for all } d \in D_{NE}, \tag{5}$$

$$X_{tv} \in \{0,1\} \qquad \text{for all } t \in T,\ v \in V, \tag{6}$$

$$U_{dt} \in \{0,1\} \qquad \text{for all } d \in D,\ t \in T, \tag{7}$$

$$Z_d^P,\ Z_d^R \in \mathbf{Z}^+ \cup \{0\} \qquad \text{for all } d \in D_{NE}, \tag{8}$$

where sets T, D, D_{NE} and V, parameters $\{n_d\}$, $\{I_{vd}\}$, $\{P_{djt}\}$, $\{Q_{dt}\}$, $\{R_{dt}\}$, and $\{m_{dt}\}$, and variables $\{X_{tv}\}$ and $\{U_{dt}\}$ are defined in VFSREP.

The objective function (O) minimizes the total weighted amount of extraimmunization of the vaccine formulary subject to the dose requirements for each disease $d \in D$ and extraimmunization restrictions for each disease $d \in D_{NE}$. The objective function weights are subjective, and hence, allow the model user to weight extraimmunization differently for each disease $d \in D_{NE}$ and/or for time periods when vaccination is permitted versus when vaccination is restricted. For example, a pediatrician may weigh more heavily those diseases that pose biological risks from extraimmunization. In the non-weighted case (i.e., $\rho_d = \gamma_d = 1$ for all disease $d \in D_{NE}$), the objective function minimizes the total number of extra doses administered for all diseases $d \in D_{NE}$. In any case, if the minimum total weighted amount of extraimmunization equals zero, then the answer to VFSREP is "yes." Constraint (1) ensures that for each disease $d \in D$, at least one vaccine that provides immunization for disease $d \in D$ is administered in some time period when dose $j = 1,2,\ldots,n_d$ may be administered. Constraint (2) and (3) guarantees that for each disease $d \in D$, at least m_{dt} doses of a vaccine that immunize against disease $d \in D$ are administered in the first $t \in T$ time periods, while also ensuring that at most one dose requirement for disease $d \in D$ is satisfied in time period $t \in T$. Constraint (4) and (5) are for each disease $d \in D_{NE}$. Constraint (4) ensures that the total number of doses administered in time periods when vaccination is permitted equals the dose requirement n_d, plus any extra doses that are administered in the time periods when vaccination is permitted. Constraint (5) ensures that the number of doses administered in time periods when vaccination is restricted equals zero, plus any extra doses that are administered in the time periods when vaccination is restricted. Constraint (6), (7), and (8) are the binary and integer constraints on the respective decision variables.

Example 8

An example of the model parameters and formulations are now given for the childhood immunization schedule in Figure 5.

<table>
<tr><th rowspan="2">DISEASE</th><th colspan="8">TIME PERIOD</th></tr>
<tr><th>1</th><th>2</th><th>3</th><th>4</th><th>5</th><th>6</th><th>7</th><th>8</th></tr>
<tr><td>1</td><td></td><td></td><td>Dose 1</td><td>Dose 2</td><td>Dose 3</td><td></td><td colspan="2">Dose 4</td></tr>
<tr><td>2</td><td></td><td></td><td>Dose 1</td><td>Dose 2</td><td colspan="4">Dose 3</td></tr>
<tr><td>3</td><td></td><td></td><td></td><td></td><td></td><td colspan="3">Dose 1</td></tr>
</table>

Figure 5: Childhood Immunization Schedule for Example 8

Here, $T = \{1,2,3,4,5,6,7,8\}$, $D = \{1,2,3\}$, and the dose vector $n = (4,3,1)$, where the d^{th} component of n corresponds to the dose requirement for disease $d = 1,2,3$. Furthermore, if $D_{NE} = \{1\}$, then the binary schedule parameters P_{djt}, Q_{dt}, and R_{dt} are:

for disease $d = 1$,

dose $j = 1$: $P_{djt} = 1(0)$ for time period $t = 3(1,2,4,5,6,7,8)$, dose $j = 2$: $P_{djt} = 1(0)$ for time period $t = 4(1,2,3,5,6,7,8)$,

dose $j = 3$; $P_{djt} = 1(0)$ for time period $t = 5(1,2,3,4,6,7,8)$, and,

dose $j = 4$; $P_{djt} = 1(0)$ for time period $t = 7,8(1,2,3,4,5,6)$, and hence,

$Q_{dt} = 1(0)$ for time period $t = 3,4,5,7,8(1,2,6)$, which implies

$R_{dt} = 1(0)$ for time period $t = 1,2,6(3,4,5,7,8)$, and

for disease $d = 2$,

dose $j = 1$: $P_{djt} = 1(0)$ for time period $t = 3(1,2,4,5,6,7,8)$,

dose $j = 2$: $P_{djt} = 1(0)$ for time period $t = 4(1,2,3,5,6,7,8)$, and

dose $j = 3$: $P_{djt} = 1(0)$ for time period $t = 5,6,7,8(1,2,3,4)$, and hence,

$Q_{dt} = 1(0)$ for time period $t = 3,4,5,6,7,8(1,2)$, and

for disease $d = 3$,

dose $j = 1$: $P_{djt} = 1(0)$ for time period $t = 6,7,8(1,2,3,4,5)$, and hence,

$Q_{dt} = 1(0)$ for time period $t = 6,7,8(1,2,3,4,5)$.

The minimum dose vectors for each disease $d \in D$ are $m_1 = (0,0,1,2,3,3,3,4)$, $m_2 = (0,0,1,2,2,2,2,3)$, and $m_3 = (0,0,0,0,0,0,0,1)$, where m_{dt} is the t^{th} component, $t = 1,2,\ldots,8$, of vector m_d for disease $d = 1,2,3$.

Suppose $V = \{1 = \{1,2,3\}\}$, which implies the binary parameters I_{vd}: $I_{1d} = 1$ for all diseases $d = 1,2,3$.

Therefore, VFSREP asks: do there exist values for the binary variables X_{tv}, $t \in T$, $v \in V$, where $X_{tv} = 1$ if vaccine $v \in V$ is administered in time period $t \in T$, 0 otherwise, and for binary variables U_{dt}, $d \in D$, $t \in T$, where $U_{dt} = 1$ if any vaccine $v \in V$ that immunizes against disease $d \in D$ is administered in time period $t \in T$, 0 otherwise, such that:

for disease $d = 1$,

dose $j = 1$: $X_{31} \geq 1$,

dose $j = 2$: $X_{41} \geq 1$,

dose $j = 3$: $X_{51} \geq 1$,

dose $j = 4$; $X_{71} + X_{81} \geq 1$, and time period

$t = 1$: $Q_{11} = 0 \Rightarrow$ no constraint,

$t = 2$: $Q_{12} = 0 \Rightarrow$ no constraint,

$t = 3$: $U_{13} \geq 1$,

t = 4: $U_{13} + U_{14} \geq 2$,

t = 5: $U_{13} + U_{14} + U_{15} \geq 3$,

t = 6: $U_{13} + U_{14} + U_{15} \geq 3$,

t = 7: $U_{13} + U_{14} + U_{15} + U_{17} \geq 3$,

t = 8: $U_{13} + U_{14} + U_{15} + U_{17} + U_{18} \geq 4$

t = 1: $Q_{11} = 0 \Rightarrow$ no constraint,

t = 2: $Q_{12} = 0 \Rightarrow$ no constraint,

t = 3: $X_{31} \geq U_{13}$,

t = 4: $X_{41} \geq U_{14}$,

t = 5: $X_{51} \geq U_{15}$,

t = 6: $X_{61} \geq U_{16}$,

t = 7: $X_{71} \geq U_{17}$,

t = 8: $X_{81} \geq U_{18}$,

Total Doses: $X_{31} + X_{41} + X_{51} + X_{71} + X_{81} = 4$,

Restricted Doses: $X_{11} + X_{21} + X_{61} = 0$, and

for disease d = 2,

dose j = 1: $X_{31} \geq 1$,

dose j = 2: $X_{41} \geq 1$,

dose j = 3: $X_{51} + X_{61} + X_{71} + X_{81} \geq 1$, and time period

t = 1: $Q_{21} = 0 \Rightarrow$ no constraint,

t = 2: $Q_{22} = 0 \Rightarrow$ no constraint,

t = 3: $U_{23} \geq 1$,

t = 4: $U_{23} + U_{24} \geq 2$,

t = 5: $U_{23} + U_{24} + U_{25} \geq 2$,

t = 6: $U_{23} + U_{24} + U_{25} + U_{26} \geq 2$,

t = 7: $U_{23} + U_{24} + U_{25} + U_{26} + U_{27} \geq 2$,

t = 8: $U_{23} + U_{24} + U_{25} + U_{26} + U_{27} + U_{28} \geq 3$

t = 1: $Q_{21} = 0 \Rightarrow$ no constraint,

t = 2: $Q_{22} = 0 \Rightarrow$ no constraint,

t = 3: $X_{31} \geq U_{23}$,

t = 4: $X_{41} \geq U_{24}$,

t = 5: $X_{51} \geq U_{25}$,

t = 6: $X_{61} \geq U_{26}$,

t = 7: $X_{71} \geq U_{27}$,

t = 8: $X_{81} \geq U_{28}$, and

for disease d = 3,

dose j = 1: $X_{61} + X_{71} + X_{81} \geq 1$, and time period

t = 1: $Q_{31} = 0 \Rightarrow$ no constraint,

t = 2: $Q_{32} = 0 \Rightarrow$ no constraint,

t = 3: $Q_{33} = 0 \Rightarrow$ no constraint,

t = 4: $Q_{34} = 0 \Rightarrow$ no constraint,

t = 5: $Q_{35} = 0 \Rightarrow$ no constraint,

t = 6: $U_{36} \geq 0$,

t = 7: $U_{36} + U_{37} \geq 0$,

t = 8: $U_{36} + U_{37} + U_{38} \geq 1$

t = 1: $Q_{31} = 0 \Rightarrow$ no constraint,

t = 2: $Q_{32} = 0 \Rightarrow$ no constraint,

t = 3: $Q_{33} = 0 \Rightarrow$ no constraint,

t = 4: $Q_{34} = 0 \Rightarrow$ no constraint,

t = 5: $Q_{35} = 0 \Rightarrow$ no constraint,

t = 6: $X_{61} \geq U_{36}$,

t = 7: $X_{71} \geq U_{37}$,

t = 8: $X_{81} \geq U_{38}$?

Therefore, if $\rho_d = \gamma_d = 1$ for disease d = 1, then the formulation for VFSREP(O) (excluding redundant constraints) for this example is:

Minimize $\quad Z_1^P + Z_1^R$

Subject to

$$X_{31} \geq 1$$

$$X_{41} \geq 1$$

$$X_{51} \geq 1$$

$$X_{71} + X_{81} \geq 1$$

$$X_{31} + X_{41} + X_{51} + X_{71} + X_{81} - Z_1^P = 4$$

$$X_{11} + X_{21} + X_{61} - Z_1^R = 0$$

$$X_{51} + X_{61} + X_{71} + X_{81} \geq 1$$

$$X_{61} + X_{71} + X_{81} \geq 1$$

$$X_{tv} \in \{0,1\} \qquad \text{for all } t \in T, v \in V$$

$$Z_1^P, Z_1^R \in \mathbf{Z}^+ \cup \{0\}. \quad \square$$

Observe that in Example 8 the constraints $\Sigma_{t=1,2,\ldots,t'} Q_{dt} U_{dt} \geq m_{dt'}$ and $\Sigma_{v\in V} Q_{dt} X_{tv} I_{vd} \geq U_{dt}$ for each disease $d \in D$ and time period $t \in T$ are redundant since each disease $d \in D$ has mutually exclusive doses. Therefore, to simplify the formulation of VFSREP(O), recall that $T_{dj} = \{t \in T : P_{djt} = 1\}$ is the set of time periods when dose $j = 1,2,\ldots,n_d$, may be administered for disease $d \in D$, where, by assumption, the time periods in T_{dj} are consecutive for all diseases $d \in D$ and doses $j = 1,2,\ldots, n_d$,. Furthermore, recall that a disease $d \in D$ is defined to have *mutually exclusive doses* if $T_{di} \cap T_{dj} = \emptyset$ for all $i, j = 1,2,\ldots,n_d$, $i \neq j$ (i.e., the sets T_{dj}, $j = 1,2,\ldots,n_d$ are pairwise mutually exclusive). Note that constraints (2) and (3) are redundant for any disease $d \in D$ with mutually exclusive doses. Furthermore, define the variable $Z_{dj}^P \in \mathbf{Z}^+ \cup \{0\}$ as the number of extra vaccine doses administered for disease $d \in D_{NE}$ in all time periods $t \in T$ such that $P_{djt} = 1$, then constraint (4) is also redundant provided that the inequality in constraint (1) becomes an equality for all disease $d \in D_{NE}$ by subtracting the slack variable Z_{dj}^P. Therefore, if every disease has mutually exclusive doses, then VFSREP(O) simplifies to the following integer program VFSREP(O)-MED (to denote the optimization model where each disease $d \in D$ has mutually exclusive doses).

VFSREP(O)-MED

Minimize $\quad \sum_{d\in D_{NE}} \left(\rho_d \left(\sum_{j=1}^{n_d} Z_{dj}^P\right) + \gamma_d Z_d^R\right)$

Subject to

$$\sum_{t\in T} \sum_{v\in V} P_{djt} X_{tv} I_{vd} \geq 1 \qquad \text{for all } d \in D_E, j = 1,2,\ldots,n_d,$$

$$\sum_{t\in T} \sum_{v\in V} P_{djt} X_{tv} I_{vd} - Z_{dj}^P = 1 \qquad \text{for all } d \in D_{NE}, j = 1,2,\ldots,n_d,$$

$$\sum_{t\in T}\sum_{v\in V} R_{dt} X_{tv} I_{vd} - Z_d^R = 0 \qquad \text{for all } d \in D_{NE},$$

$$X_{tv} \in \{0,1\} \qquad \text{for all } t \in T,\ v \in V,$$

$$Z_{dj}^P \in \mathbf{Z}^+ \cup \{0\} \qquad \text{for all } d \in D_{NE},\ j = 1,2,\ldots,n_d,$$

$$Z_d^R \in \mathbf{Z}^+ \cup \{0\} \qquad \text{for all } d \in D_{NE}.$$

All of the diseases in the 2006 Recommended Childhood Immunization Schedule have mutually exclusive doses, though some diseases in past schedules did not have this property. For example, hepatitis B did not have mutually exclusive doses in the 2005 Recommended Childhood Immunization Schedule (CDC 2005). Therefore, the simplification of VFSREP(O) to VFSREP(O)-MED has practical implications.

There is an alternative formulation of VFSREP(O)-MED that has both theoretical and practical value. To describe this formulation, define the binary decision variable $\pi_{dj} = 1$ if the j^{th} dose requirement for disease $d \in D$ is satisfied, 0 otherwise (i.e., $\pi_{dj} = 1$ if and only if $\Sigma_{t\in T}\Sigma_{v\in V} P_{djt} X_{tv} I_{vd} \geq 1$ for $d \in D_E$, $j = 1,2,\ldots,n_d$ and $\Sigma_{t\in T}\Sigma_{v\in V} P_{djt} X_{tv} I_{vd} = 1$ for $d \in D_{NE}$, $j = 1,2,\ldots,n_d$). Therefore, an alternative formulation for VFSREP-MED (i.e., VFSREP(O)-MED(A), where (A) denotes alternative) is the following integer program.

VFSREP(O)-MED(A)

Maximize $\sum_{d\in D}\sum_{j=1}^{n_d} \pi_{dj}$

Subject to

$$\sum_{t\in T}\sum_{v\in V} P_{djt} X_{tv} I_{vd} \geq \pi_{dj} \qquad \text{for all } d \in D_E,\ j = 1,2,\ldots,n_d,$$

$$\sum_{t\in T}\sum_{v\in V} P_{djt} X_{tv} I_{vd} = \pi_{dj} \qquad \text{for all } d \in D_{NE},\ j = 1,2,\ldots,n_d,$$

$$\sum_{t\in T}\sum_{v\in V} R_{dt} X_{tv} I_{vd} = 0 \qquad \text{for all } d \in D_{NE},$$

$$X_{tv} \in \{0,1\} \qquad \text{for all } t \in T,\ v \in V,$$

$$\pi_{dj} \in \{0,1\} \qquad \text{for all } d \in D,\ j = 1,2,\ldots,n_d.$$

Note that the objective function of VFSREP(O)-MED(A) is maximized when $\pi_{dj} = 1$ for all diseases $d \in D$ and doses $j = 1,2,\ldots,n_d$, and hence, if the optimal solution equals $\Sigma_{d=1}^{\delta} n_d$, then the j^{th} dose requirement for every disease $d \in D$ is satisfied and no extra vaccine doses were administered for any disease $d \in D_{NE}$. Therefore, VFSREP(O)-MED(A) maximizes the number of doses that may be administered while forbidding extraimmunization. Furthermore, all the decision variables in VFSREP(O)-MED(A) are binary, which offers theoretical advantages (see Nemhauser and Wolsey 1999). For example, VFSREP(O)-MED(A) shares a similar structure to the MAX-SAT problem (Hochbaum 1997) or, when $\delta_{NE} = \delta$, to the Set-Partitioning problem (Nemhauser and Wolsey 1999), both of which have been well-studied.

Moreover, assigning weights to the decision variables in VFSREP(O)-MED(A) offers another practical application. Each vaccine has a known *immunogenicity* (the ability of the vaccine to immunize against a disease). For example, the immunogenicity of a Hepatitis B vaccine is >95%, which means at least 95% of those children that receive the three dose series develop a protective antibody response against the disease (CDC 2003). However, for some vaccines, the immunogenicity increases with each dose. For example, for a Hepatitis B vaccine the immunogenicity is 30-50% after the first dose, 75% after the second, and >95% after the third dose (CDC 2003). Therefore, if each objective function decision variable π_{dj} in VFSREP(O)-MED(A) were weighted by immunogenicity, then VFSREP(O)-MED(A) would maximize the total immunogenicity of the vaccine formulary while restricting extraimmunization. Specifically, define Imm_{dj} as the immunogenicity for disease $d \in D$ after dose $j = 1,2,\ldots,n_d$ is administered, and let $\omega_{dj} = 1 - Imm_{dj}$ be the corresponding objective function coefficient for decision variable π_{dj}. This assumes equivalent immunogenicities for each vaccine $v \in V$ that immunizes against disease $d \in D$. Weighting the decision variables accordingly would require the additional constraint $\pi_{d1} \leq \pi_{d2} \leq \cdots \leq \pi_{dn_d}$ for all diseases $d \in D$ in the formulation of VFSREP(O)-MED(A).

4.2 Computational Complexity

This section presents the computational complexity of VFSREP and VFSREP(O). In the worst case, these problems are shown to be intractable. There are, however, some special cases that are solvable in polynomial time. Theorem 7 states that VFSREP is *NP*-complete.

THEOREM 7: *VFSREP is NP-complete in the strong sense.*

PROOF: First, VFSREP is in *NP* since given a set of guessed values for the binary variables X_{tv}, $t \in T$, $v \in V$, and U_{dt}, $d \in D$, $t \in T$, $\Sigma_{t\in T}\Sigma_{v\in V} P_{djt} X_{tv} I_{vd} \geq 1$ for all $d \in D$, $j = 1,2,\ldots,n_d$, $\Sigma_{t=1,2,\ldots,t'} Q_{dt} U_{dt} \geq m_{dt'}$ for all $d \in D$ and $t' \in T$, and $\Sigma_{v\in V} Q_{dt} X_{tv} I_{vd} \geq U_{dt}$ for all $d \in D$ and $t \in$

T can all be checked in $O(\upsilon\tau^2\delta)$ time, and the constraints $\Sigma_{t\in T}\Sigma_{v\in V}\, R_{dt}\, X_{tv}\, I_{vd} = 0$ and $\Sigma_{t\in T}\Sigma_{v\in V}\, Q_{dt}\, X_{tv}\, I_{vd} = n_d$ for all $d \in D_{NE}$ can be checked in $O(\upsilon\tau\delta_{NE})$ time.

To show that VFSREP is *NP*-complete, a polynomial transformation from 1-in-3 3-SAT with 2-SAT to VFSREP can be constructed.

Given an arbitrary instance of 1-in-3 3-SAT with 2-SAT, define a particular instance of VFSREP as follows: Set $T = \{1\}$; $D = D_{NE} = \{1,2,\ldots,m+n\}$, $\delta = m+n$; $D_E = \emptyset$; $V = \{1,2,\ldots,2n\}$, $\upsilon = 2n$; and $n_1 = n_2 = \cdots = n_{m+n} = 1$. Let $y_1,y_2,\ldots,y_n$ correspond to vaccines $v = 1,2,\ldots,n$, respectively, and $1\text{-}y_1,1\text{-}y_2,\ldots,1\text{-}y_n$ correspond to vaccines $v = n+1,n+2,\ldots,2n$, respectively. Let clauses $C_1,C_2,\ldots,C_m$ correspond to diseases $d = 1,2,\ldots,m$, respectively, and $C_{m+1},C_{m+2},\ldots,C_{m+n}$ correspond to diseases $d = m+1,m+2,\ldots,m+n$, respectively. Set the binary parameters as follows:

$$I_{vd} = \begin{cases} 1 & \text{if the literal } y_v \text{ is in clause } \mathrm{C}_d \text{ for } v = 1,2,\ldots,n;\, d = 1,2,\ldots,m+n, \text{respectively} \\ 0 & \text{otherwise,} \end{cases}$$

$$I_{(n+v)d} = \begin{cases} 1 & \text{if literal } (1 - y_v) \text{ is in clause } \mathrm{C}_d \text{ for } v = 1,2,\ldots,n;\, d = 1,2,\ldots,m+n, \text{respectively} \\ 0 & \text{otherwise.} \end{cases}.$$

Therefore, the vaccines that immunize against disease $d = 1,2,\ldots,m+n$ are determined by the literals in clause C_d. Set $P_{djt} = 1$ for $d \in D$, $j = 1$, and $t = 1$, and $Q_{dt} = 1$ for $d \in D$, $t = 1$. Lastly, $R_{dt} = 0$ for $d \in D$, $t = 1$. Clearly, this transformation can be made in polynomial time in the size of the arbitrary instance of 1-in-3 3-SAT with 2-SAT. Furthermore, this transformation results in a particular instance of VFSREP where each $d \in D$ has mutually exclusive doses, and since $D_E = \emptyset$, the only constraints for this particular instance are $\Sigma_{t\in T}\Sigma_{v\in V}\, P_{djt}\, X_{tv}\, I_{vd} = 1$ for all $d \in D$, or simply, $\Sigma_{v\in V}\, X_{1v}\, I_{vd} = 1$ for all $d \in D$.

To complete the proof, it is necessary to show that a *yes* for this particular instance of VFSREP implies a *yes* for the arbitrary instance of 1-in-3 3-SAT with 2-SAT, and a *yes* for the arbitrary instance of 1-in-3 3-SAT with 2-SAT implies a *yes* for this particular instance of VFSREP.

Suppose the answer to the particular instance of VFSREP is *yes*. Then there exist values for the binary variables X_{1v}, $v \in V$, such that $\Sigma_{v\in V}\, X_{1v}\, I_{vd} = 1$ for all $d \in D$. Clearly, $I_{vd} = 1$ for $v \in V$, $d \in D$, corresponds to a literal (y_j or $1- y_j$ for some $j = 1,2,\ldots,n$) that is in clause C_i, $i = 1,2,\ldots,m+n$. Therefore, if $\Sigma_{v\in V}\, X_{1v}\, I_{vd} = 1$ for all $d \in D$, then the binary variable with $X_{1v} = 1$

for each constraint corresponds to the one literal that satisfies clause C_i, for $i = 1,2,\ldots,m+n$. Moreover, $\Sigma_{v \in V}\ X_{1v} I_{vd} = 1$ for all $d = m+1, m+2, \ldots, m+n$, and since vaccines $v = k$ and $v = n+k$ immunize against disease $d = m+k$, then the binary variables X_{1k} and $X_{1(n+k)}$ exist together in the constraint for disease $d = m+k$. This shows that both y_j and $1\text{-}y_j$ cannot be one for all $j = 1,2,\ldots,n$, which means there is a Boolean variable assignment that satisfies all $m+n$ clauses with exactly one true literal, and hence, the answer to the arbitrary instance of 1-in-3 3-SAT with 2-SAT is *yes*.

Now suppose the answer to the arbitrary instance of 1-in-3 3-SAT with 2-SAT is *yes*. Then there exists a Boolean variable assignment that results in all $m+n$ clauses being satisfied by exactly one literal. For each Boolean variable y_j, $j = 1,2,\ldots,n$, there are two corresponding binary variables where one such variable (X_{1j}) corresponds to y_j and the other variable ($X_{1(n+j)}$) corresponds to $1 - y_j$. Therefore, if $y_j = 1(0)$, set $X_{1j} = 1(0)$ and $X_{1(n+j)} = 0(1)$ for $j = 1,2,\ldots,n$. The claim is that these values for X_{1v}, $v = 1,2,\ldots,2n$, result in a *yes* answer for the particular instance of VFSREP. Suppose not, that is, suppose there does not exist values for the binary variables X_{1v}, $v = 1,2,\ldots,2n$, such that $\Sigma_{v \in V}\ X_{1v} I_{vd} = 1$ for all $d \in D$. As seen above, the constraints for $d = m+1, m+2, \ldots, m+n$ correspond to the j^{th} literal pair y_j and $1\text{-}y_j$ for $j = 1,2,\ldots,n$, and hence, if these constraints are not satisfied then neither y_j nor $1\text{-}y_j$ equal one, which is a contradiction. Therefore, for all possible binary variable values of X_{1v}, $v = 1,2,\ldots,2n$, there must exist some $d = 1,2,\ldots,m$ such that $\Sigma_{v \in V}\ X_{1v} I_{vd} \neq 1$, and hence, $\Sigma_{v \in V}\ X_{1v} I_{vd} = 0$ or $\Sigma_{v \in V}\ X_{1v} I_{vd} > 1$. In either case, the binary variable values of X_{1v} correspond to the Boolean variable values for y_j, $j = 1,2,\ldots,n$, and I_{vi} identifies the literals in clause C_i for $v = 1,2,\ldots,2n$ and $i = 1,2,\ldots,m$. Therefore, if for all possible binary variable values, $\Sigma_{v \in V}\ X_{1v} I_{vd} = 0$ for some disease $d = i$, then clause C_i is not satisfied, which contradicts 1-in-3 3-SAT with 2-SAT being *yes*. Likewise, $\Sigma_{v \in V}\ X_{1v} I_{vd} > 1$ for some disease $d = i$ implies clause C_i is satisfied by more than one literal, which, again, is a contradiction. Therefore, the values for X_{1v}, $v = 1,2,\ldots,2n$, defined above result in a *yes* answer for the particular instance of VFSREP.

Furthermore, note that VFSREP is not a number problem (see Garey and Johnson 1979) since the only numbers occurring in an instance of VFSREP are the dose requirements n_d for all $d \in D$, which are clearly bounded by $\tau = |T|$, and hence, by the length of the instance. Therefore, VFSREP is strongly *NP*-complete. ■

The proof of Theorem 7 suggests several special cases of VFSREP that remain *NP*-complete. In particular, VFSREP remains *NP*-complete when there exists only one time period (i.e., $\tau = 1$), when extraimmunization is restricted for all diseases $d \in D$ (i.e., $D_E = \emptyset$), and when each disease requires only one dose of vaccine (i.e., $n_d = 1$ for all diseases $d \in D$). Theorem 8 gives some additional special cases of VFSREP that remain *NP*-complete.

THEOREM 8: *The following special cases of VFSREP are NP-complete in the strong sense*:

i) *Only one vaccine exists* (*i.e.*, $\upsilon = 1$ *where* $I_{vd} = 1$ *for all diseases* $d \in D$),
ii) *The disease set has cardinality of at least three* (*i.e.*, $\delta \geq 3$),
iii) *Every vaccine is at least a trivalent vaccine* (*i.e.*, $Val(v) \geq 3$ *for all vaccines* $v \in V$)
iv) *Extraimmunization is restricted for only one disease* (*i.e.*, $\delta_{NE} = 1$).

PROOF: To show *i*), *ii*), and *iii*) a polynomial transformation from 3DM to VFSREP is constructed, where $\delta = 3$ and $V = \{v\}$ is a trivalent vaccine

Given an arbitrary instance of 3DM define a particular instance of VFSREP as follows: Set $T = M$, $D = D_{NE} = \{1,2,3\}$, $D_E = \emptyset$, $V = \{v\}$, and $n_1 = n_2 = n_3 = q$. Let the q elements in W, Y, and Z correspond to the doses of disease $d = 1, 2, 3$, respectively. Therefore, w_1 corresponds to the first dose of vaccine for disease $d = 1$, w_2 corresponds to the second dose of vaccine for disease $d = 1$, and so forth through dose q. Furthermore, since $T = M$, then the 3-tuple m_i corresponds to the i^{th} time period. Set the binary parameters $I_{vd} = 1$ for all $d \in D$. Set $P_{djt} = 1$, $d = 1$, $j = 1,2,\ldots q$, $t = i$, $i = 1,2,\ldots,k$, if element $w_j \in m_i$, 0 otherwise; likewise, $P_{djt} = 1$, $d = 2$, $j = 1,2,\ldots q$, $t = i$, $i = 1,2,\ldots,k$, if element $y_j \in m_i$, 0 otherwise; and $P_{djt} = 1$, $d = 3$, $j = 1,2,\ldots q$, $t = i$, $i = 1,2,\ldots,k$, if element $z_j \in m_i$, 0 otherwise. Set $Q_{dt} = 1$ and $R_{dt} = 0$ for all $d \in D$, $t \in T$ since some dose of disease $d \in D$ is permitted in time $t \in T$, and lastly, set $m_{dt} = 0$ for all $d \in D$, $t = 1,2,\ldots,k$-1, and $m_{dt} = q$ for $d \in D$, $t = k$. Clearly, this transformation can be done in polynomial time in the size of the arbitrary instance of 3DM.

To complete the proof, it is necessary to show that a *yes* for the particular instance of VFSREP implies a *yes* for the arbitrary instance 3DM and a *yes* for the arbitrary instance 3DM implies a *yes* for the particular instance of VFSREP.

Suppose the answer to the particular instance of VFSREP is *yes*. Then there exist values for the binary variables X_{t1}, $t = 1,2,\ldots,|M|$, and U_{dt}, $d \in D$, $t \in T$, such that

$$\sum_{t\in T} P_{djt} X_{t1} \geq 1 \quad \text{for all } d \in D, j = 1,2,..,q; \tag{1}$$

$$\sum_{t\in T} Q_{dt} U_{dt} \geq q \quad \text{for all } d \in D; \tag{2}$$

$$X_{t1} \geq U_{dt} \qquad \text{for all } d \in D, t \in T; \tag{3}$$

$$\sum_{t\in T} Q_{dt} X_{t1} = q \text{ (since } n_d = q \text{ for all } d \in D). \tag{4}$$

Observe that in any time period $t \in T$, exactly one dose for each disease $d \in D$ is permitted, that is, given d and t, $P_{djt} = 1$ for some $j = 1,2,\ldots,q$. Therefore, constraints (2) and (3) imply $\Sigma_{t\in T}\, Q_{dt}\, X_{t1} \geq q$ for all $d \in D$ and becomes the single constraint $\Sigma_{t\in T}\, X_{t1} \geq q$. This observation along with constraint (4) (i.e., $\Sigma_{t\in T}\, Q_{dt}\, X_{t1} = q$) implies $\Sigma_{t\in T}\, X_{t1} = q$, which also means constraints (1) are tight (i.e., $\Sigma_{t\in T}\, P_{djt}\, X_{t1} = 1$ for all $d \in D$, $j = 1,2,..,q$) or some dose would not be satisfied. Therefore, the constraints for this particular instance of VFSREP become

$$\sum_{t\in T} P_{djt} X_{t1} = 1 \text{ for all } d \in D, j = 1,2,..,q; \tag{1'}$$

$$\sum_{t\in T} X_{t1} = q. \tag{2'}$$

The claim is that if $X_{tv} = 1$, then the 3-tuple $m \in M$ corresponding to time period $t \in T$ is part of the matching $M' \subseteq M$. Let M' be the set of 3-tuples in M corresponding to $X_{tv} = 1$. Constraints (1') ensure there exists exactly one time period $t \in T$ that vaccine v is administered to satisfy each dose requirement for every $d \in D$. Since the j^{th} dose corresponds to some element in $w_j \in W$, $y_j \in Y$, or $z_j \in Z$, then exactly one $m \in M'$ contains element w_j, y_j, and z_j. This means no two elements of M' agree in any coordinate. Constraint (2') ensures that vaccine v is administered in exactly q time periods, and hence, $|M'| = q$. Therefore, the answer to the arbitrary instance of 3DM is *yes*.

Now suppose the answer to the arbitrary instance of 3DM is *yes*. Then there exists a matching $M' \subseteq M$ such that $|M'| = q$ and no two elements of M' agree in any coordinate. Since each 3-tuple in M for the arbitrary instance of 3DM corresponds to a different time period for the particular instance of VFSREP, set $X_{tv} = 1$ if the corresponding 3-tuple m is in M', and zero otherwise. Since M' is a matching, then each element of W, Y, and Z exists in exactly one $m \in M'$, and hence, each dose for each disease is satisfied exactly once in the time period corresponding to m (i.e., $\Sigma_{t\in T}\, P_{djt}\, X_{tv} = 1$ for all $d \in D$, $j = 1,2,\ldots,q$). Moreover, $|M'| = q$, where no two elements of M' agree in any coordinate, which implies $\Sigma_{t\in T}\, X_{tv} = q$ for all $d \in D$. Extraimmunization occurs when either more than q doses of vaccine are administered for any disease $d \in D$ or a dose of vaccine is administered in some time period where a dose is not required. Clearly, the number of doses can not exceed q for any disease $d \in D$ since vaccine v is administered in exactly q time periods, and, by construction, every time period covers some dose for each disease $d \in D$ (i.e., $R_{dt} = 0$ for all $d \in D$, $t \in T$). Therefore, no extraimmunization exists, and hence, the answer to the particular instance of VFSREP is *yes*.

To show *iv*), a polynomial transformation from 3-SAT to VFSREP is constructed, where $\delta_{NE} = 1$.

First, define n additional 2-SAT clauses within the arbitrary 3-SAT instance, specifically, clauses $C_{m+1}, C_{m+2},\ldots,C_{m+n}$, where clause $C_{m+k} = (y_k \vee (1\text{-}y_k))$ for $k = 1,2,\ldots,n$. Clearly, this does not change the solution of the 3-SAT instance since each Boolean variable is either true (1) or false (0), and this transformation can be done in linear time. For VFSREP, Set $T = \{1,2,\ldots,2n\}$, $\tau = 2n$; $D_E = \{1,2,\ldots,m\}$; $D_{NE} = \{m+1\}$ (i.e., $\delta_{NE} = 1$), and $\delta = m+1$; $V = \{v\}$; and $n_1 = n_2 = \cdots = n_m = 1$ and $n_{m+1} = n$ (n being the number of Boolean variables). Let $y_1,y_2,\ldots,y_n$ correspond to the odd time periods $1,3,\ldots,2n–1$, respectively, and $1\text{-}y_1,1\text{-}y_2,\ldots,1\text{-}y_n$ correspond to the even time periods $2,4,\ldots,2n$, respectively. Let clauses $C_1,C_2,\ldots,C_m$ correspond to diseases $1,2,\ldots,m$, respectively, and clause C_{m+j} correspond to dose $j = 1,2,\ldots,n$ for disease $d = m+1$, respectively. Therefore, the 3-SAT clauses correspond to the diseases where extraimmunization is permitted and the 2-SAT clauses correspond to the n doses for disease $d = m+1$ where no extraimmunization is permitted. Set the binary parameters as follows:

$$P_{d=i,j=1,t=2k-1} = \begin{cases} 1 & \text{if literal } y_k \text{ is in clause } C_i \text{ for } k = 1,2,\ldots,n; i = 1,2,\ldots,m, \text{respectively} \\ 0 & \text{otherwise,} \end{cases}$$

$$P_{d=i,j=1,t=2k} = \begin{cases} 1 & \text{if literal } (1 - y_k) \text{ is in clause } C_i \text{ for } k = 1,2,\ldots,n; i = 1,2,\ldots,m, \text{respectively} \\ 0 & \text{otherwise.} \end{cases}$$

Therefore, the times periods where the one dose of vaccine may be administered to immunize against disease $d = i$ are determined by the literals in clause C_i, $i = 1,2,\ldots,m$. Set $Q_{dt} = P_{d1t}$ for all $d \in D_E$ and $t \in T$. Furthermore, the first dose of vaccine for disease $d = m+1$ must be administered in time period $t = 1$ or $t = 2$ as determined by clause C_{m+1}, the second dose of vaccine for disease $d = m+1$ must be administered in time period $t = 3$ or $t = 4$ as determined by clause $C_{m+2,}$ etc., until the final dose (dose n) of vaccine for disease $d = m+1$, which must be administered in time period $t = 2n - 1$ or $t = 2n$, and hence, $P_{d=m+1,j=k,t=2k-1} = P_{d=m+1,j=k,t=2k} =$ 1 for $k = 1,2,\ldots,n$, and 0, otherwise. Set $I_{vd} = 1$ for all $d \in D$ and $v = 1$. Lastly, $Q_{dt} = 1$ and $R_{dt} = 0$ for $d = m+1$, $t \in T$. Clearly, this transformation can be made in polynomial time in the size of the arbitrary instance of 3-SAT. Furthermore, this transformation results in a particular instance of VFSREP where each $d \in D$ has mutually exclusive doses.

It remains to show that a *yes* for this particular instance of VFSREP implies a *yes* for the arbitrary instance of 3-SAT, and a *yes* for the arbitrary instance of 3-SAT implies a *yes* for this particular instance of VFSREP.

Suppose the answer to the particular instance of VFSREP is *yes*. Since all $d \in D$ in the particular instance of VFSREP have mutually exclusive doses, then there exist values for the binary variables X_{t1}, $t \in T$, such that $\Sigma_{t\in T} P_{djt} X_{t1} \geq 1$ for all $d \in D_E$, $j = 1$, and $\Sigma_{t\in T} P_{djt} X_{t1} = 1$ for $d = m+1$, $j = 1,2,\ldots,n$. Clearly, $P_{djt} = 1$ for $d \in D$, $j = 1,2,\ldots n$, $t \in T$, corresponds to a literal (y_k or $1\text{-}y_k$ for some $k = 1,2,\ldots,n$) that is in clause C_i, $i = 1,2,\ldots,m+n$. Therefore, if $\Sigma_{t\in T} P_{djt} X_{t1} \geq 1$ for all $d \in D_E$, then the binary variables with $X_{t1} = 1$ for each constraint corresponds to literals that satisfy clause C_i, for $i = 1,2,\ldots,m$. Similarly, if $\Sigma_{t\in T} P_{djt} X_{t1} = 1$ for $d = m+1$, $j = 1,2,\ldots,n$, then the binary variable with $X_{t1} = 1$ for each constraint corresponds to the one literal that satisfies clause C_i, for $i = m+1,m+2,\ldots,m+n$. Moreover, since the j^{th} dose of vaccine v must be administered in time period $t = 2j - 1$ or time period $t = 2j$, to immunize against disease $d = m+1$, the binary variables $X_{(2j-1)1}$ and $X_{(2j)1}$ exist together in the constraints for disease $d = m+1$, $j = 1,2,\ldots,n$. This shows that both y_k and $1\text{-}y_k$ cannot be one for all $k =$ $1,2,\ldots,n$, which means there is a Boolean variable assignment that satisfies all $m + n$ clauses, hence the answer to the arbitrary instance of 3-SAT is *yes*.

Now suppose the answer to the arbitrary instance of 3-SAT is *yes*. Then there exists a Boolean variable assignment that results in all m clauses being satisfied. For each Boolean variable y_k, $k = 1,2,\ldots,n$, there are two corresponding binary variables where one such variable ($X_{(2k-1)1}$) corresponds to y_k and the other variable ($X_{(2k)1}$) corresponds to 1- y_k. Therefore, if $y_k =1$ (0), set $X_{(2k-1)1} = 1$ (0) and $X_{(2k)1} = 0$ (1). The claim is these values for X_{t1}, $t \in T$, result in a *yes* answer for the particular instance of VFSREP. Suppose not, that is, suppose there does not exist values for the binary variables X_{t1}, $t \in T$, such that $\Sigma_{t\in T} P_{djt} X_{t1} \geq$ 1 for all $d \in D_E$ and $\Sigma_{t\in T} P_{djt} X_{t1} = 1$ for $d = m+1$, $j = 1,2,\ldots,n$. As before, the constraints $\Sigma_{t\in T}$

$P_{djt}\,X_{t1}$ = 1 for $d = m+1$, $j = 1,2,\ldots,n$ correspond to the k^{th} literal pair y_k and $1\text{-}y_k$. Therefore, if these constraints are not satisfied then neither y_k nor $1\text{-}y_k$ equal one, which is a contradiction, and hence, for all possible binary variable values of X_{t1}, $t \in T$, there must exist some $d \in D_E$ such that $\Sigma_{t\in T}\, P_{djt}\, X_{t1} = 0$ when $j = 1$. However, the binary variable values of X_{t1} correspond to the Boolean variable values for y_k, $k = 1,2,\ldots,n$, and P_{djt} identifies the literals in clause C_i for $t \in T$, $j = 1$, $i = 1,2,\ldots,m$. Therefore, if for all possible binary variable values, $\Sigma_{t\in T}\, P_{djt}\, X_{t1} = 0$ for some disease $d = i$ and $j = 1$, then clause C_i is not satisfied, which contradicts 3-SAT being *yes*. Therefore, the values for X_{t1}, $t \in T$, defined above result in a *yes* answer for the particular instance of VFSREP.

Furthermore, note that VFSREP for special cases *i*)-*iv*) are not number problems (see Garey and Johnson 1979) since the only numbers are the dose requirements n_d for all $d \in D$, which are clearly bounded by $\tau = |T|$, and hence, by the length of the instance. Therefore, these special cases of VFSREP are strongly *NP*-complete. ■

Theorems 7 and 8 imply VFSREP remains *NP*-complete even when the sets T, D, D_{NE}, and V, or when the dose (n_d, $d \in D$) parameters are significantly restricted. In addition, since VFSREP is *NP*-complete, then the corresponding optimization problem VFSREP(O) is *NP*-hard. Furthermore, all of the diseases $d \in D$ in the polynomial transformation in the proof of Theorem 7 have mutually exclusive doses, and hence, VFSREP(O)-MED is also *NP*-hard. Another facet to the complexity of VFSREP lies in the flexibility of the childhood immunization schedule. In general, VFSREP becomes more difficult if the doses for each disease may be administered in several time periods (i.e., for a given disease $d \in D$ and dose $j = 1,2,\ldots,n_d$, $P_{djt} = 1$ for multiple time periods $t \in T$). Recall that a childhood immunization schedule as *tight* if every required dose of vaccine for each disease $d \in D$ may be administered in exactly one time period (i.e., for dose $j = 1,2,\ldots,n_d$ and disease $d \in D$, $P_{djt} = 1$ for exactly one time period $t \in T$). A tight schedule implies that all diseases $d \in D$ have mutually exclusive doses, since dose $j = 1,2,\ldots,n_d$ may be administered in exactly one time period, and hence, the time period $t \in T$ when $P_{djt} = 1$ is unique. Furthermore, define a childhood immunization schedule as *aligned* if given any disease $d' \in D$ and dose $j = 1,2,\ldots,n_{d'}$, $T_{d'j} = \{t \in T: P_{d'jt} = 1\} = T_{dk} = \{t \in T: P_{dkt} = 1\}$ for every disease $d \in D$ and dose $k = 1,2,\ldots,n_d$ such that $T_{d'j} \cap T_{dk} \neq \emptyset$ (i.e., the set of time periods when dose j may be administered for disease $d' \in D$ are identical to the set of time periods when dose k may be administered for every disease $d \in D$ such that the set of time periods when dose j may be administered for disease $d' \in D$ overlap with the set of time periods when dose k may be administered for every disease $d \in D$). Moreover, a childhood immunization schedule is *completely aligned* if $n_d = N$ for all diseases $d \in D$ and $T_{d'j} = \{t \in T: P_{d'jt} = 1\} = T_{dj} = \{t \in T: P_{djt} = 1\}$ for every disease $d', d \in D$ and dose $j = 1,2,\ldots,N$. By these definitions, both tight and/or aligned childhood immunization schedules are less flexible. Figure 6 displays a tight childhood immunization schedule that is also completely aligned, and Figure 7 displays an aligned childhood immunization schedule.

DISEASE	**TIME PERIOD**					
	1	**2**	**3**	**4**	**5**	**6**
1	Dose 1	Dose 2		Dose 3		Dose 4
2	Dose 1	Dose 2		Dose 3		Dose 4
3	Dose 1	Dose 2		Dose 3		Dose 4

Figure 6: A Tight and Completely Aligned Childhood Immunization Schedule

<table>
<tr><th rowspan="2">DISEASE</th><th colspan="6">TIME PERIOD</th></tr>
<tr><th>1</th><th>2</th><th>3</th><th>4</th><th>5</th><th>6</th></tr>
<tr><td>1</td><td>Dose 1</td><td colspan="2">Dose 2</td><td></td><td colspan="2">Dose 3</td></tr>
<tr><td>2</td><td>Dose 1</td><td></td><td></td><td></td><td colspan="2">Dose 2</td></tr>
<tr><td>3</td><td></td><td colspan="2">Dose 1</td><td></td><td colspan="2">Dose 2</td></tr>
</table>

Figure 7: An Aligned Childhood Immunization Schedule

By design, VFSREP is a feasibility problem constrained by dosage requirements for all diseases $d \in D$ and restrictions on extraimmunization for all diseases $d \in D_{NE}$, and hence, when there is no restriction on extraimmunization (i.e., $D_{NE} = \emptyset$), then VFSREP is solvable in polynomial time. In fact, the trivial solution $X_{tv} = 1$ for all time periods $t \in T$ and vaccines $v \in V$ satisfies all dose requirements, assuming that there exists at least one vaccine $v \in V$ such that $I_{vd} = 1$ for all diseases $d \in D$.

Other special cases of VFSREP that are solvable in polynomial time occur when the valency of the vaccine set is limited to monovalent vaccines, the number of vaccines that immunize against each disease is limited to two vaccines, the childhood immunization schedule is tight or aligned, and when the number of diseases is less than three. To see this, first consider limitations on the valency of the vaccine set. If every vaccine $v \in V$ is monovalent, then VFSREP is solvable in $O(\tau\delta)$ time. Lemma 6 considers a stronger result for the case when all diseases $d \in D$ have a corresponding monovalent vaccine $v \in V$.

LEMMA 6: *Given any disease $d \in D$, if there exists a vaccine $v \in V$ such that $I_{vd} = 1$ and $Val(v) = 1$, then VFSREP is solvable in $O(\tau\delta)$ time.*

PROOF: The feasible vaccine schedule that restricts extraimmunization for all diseases $d \in D_{NE}$ may be found by looping through the set of time periods and diseases and administering the corresponding monovalent vaccine for dose $j = 1,2, \ldots, n_d$, in the first time period when $P_{djt} = 1$. ■

Define the linear programming (LP) relaxation of VFSREP(O)-MED as the LP model of VFSREP(O)-MED along with the relaxed binary variable constraint $0 \leq X_{tv} \leq 1$ for all time periods $t \in T$ and vaccines $v \in V$ and, for all diseases $d \in D_{NE}$, the relaxed integer variable constraints $Z_{dj}^{P} \geq 0$ for dose $j = 1,2,\ldots,n_d$, and $Z_{d}^{P} \geq 0$. Theorem 9 states a similar result for VFSREP(O)-REP to Lemma 6 for the case when all vaccines $v \in V$ are monovalents.

THEOREM 9: *If $Val(v) = 1$ for all vaccines $v \in V$, then the LP relaxation of VFSREP(O)-MED yields an optimal integer solution.*

PROOF: This results assumes that the time periods where $P_{djt} = 1$ for all diseases $d \in D$, and doses $j = 1,2,\ldots,n_d$, are consecutive. This assumption holds in the 2006 Recommended Childhood Immunization Schedule, as well as all the randomly generated childhood immunization schedules presented in Section 4.4.

Suppose for a given instance of VFSREP(O)-MED that $Val(v) = 1$ for all vaccines $v \in V$. Consider the LP relaxation of VFSREP(O)-MED and denote its constraint matrix by **A**. If **A** is totally unimodular, then every basic feasible solution is integer, provided the right-hand-side (rhs) vector is integer (Ahuja et al. 1993). Clearly, the rhs vector in the LP relaxation of VFSREP(O)-MED is integer, and hence, it remains to show that **A** is indeed totally unimodular.

By definition, **A** is totally unimodular if every square submatrix of **A** has determinant 0, 1, or -1. It is well known that **A** is totally unimodular if the non-zero elements in each row are in consecutive columns (known as the *consecutive ones* property). Without loss of generality, assume that there is exactly one vaccine $v \in V$ that provides immunization against each disease $d \in D$. Furthermore, it is sufficient to consider only the columns of **A** corresponding to the decision variables X_{tv}, $t \in T$, $v \in V$, since the remaining columns correspond to slack variables, which have exactly one non-zero row, and hence, may be ordered at will within the matrix **A**. Moreover, assume the columns of **A** are ordered according to the set V. For example, the first τ columns of **A** correspond to the decision variables associated with the first vaccine, (i.e., column t corresponds to decision variable X_{t1}), the second τ columns of **A** correspond to the decision variables associated with the second vaccine, and so on.

Now consider some disease $d \in D$, and let $v \in V$ be a vaccine such that $I_{vd} = 1$. Since $Val(v) = 1$, then the only non-zero entries in any row of **A** that corresponds to disease $d \in D$ must be in the τ consecutive columns corresponding to vaccine $v \in V$. By assumption, the time periods when $P_{djt} = 1$ for disease $d \in D$, dose $j = 1,2,\ldots,n_d$, are consecutive, and hence, the rows corresponding to each dose for disease $d \in D$ have the consecutive ones property. Furthermore, if $d \in D_{NE}$ and if there exists some time period $t \in T$ such that $P_{djt} = 0$ for all $j = 1,2,\ldots,n_d$, which means that no dose for disease $d \in D$ is permitted in time period $t \in T$, then $R_{dt} = 1$. Therefore, every row of **A** has a one in the column corresponding to the variable X_{tv}, where $R_{dt} = 1$, and hence, the columns corresponding to these variables may be reordered so that the consecutive ones property is satisfied. Therefore, **A** is totally unimodular, since the non-zero elements in each row of **A** are in consecutive columns. ■

Theorem 9 also implies that VFSREP(O)-MED is solvable in polynomial time when all vaccines $v \in V$ are monovalents, since linear programming is solvable in polynomial time (Bazaraa et al. 1990). Moreover, Theorem 9 may be used to show that the heuristics presented in Section 4.3 return the optimal solution when all vaccines $v \in V$ are monovalents.

Given a tight, aligned, or completely aligned childhood immunization schedule and a vaccine set where there are at most two vaccines that immunize against each disease $d \in D$ (i.e., $\Sigma_{v\in V} I_{vd} \leq 2$ for all diseases $d \in D$), Lemmas 7 and 8 yield other special cases of VFSREP that are solvable in polynomial time.

LEMMA 7: *Given a tight childhood immunization schedule, if $\Sigma_{v\in V} I_{vd} \leq 2$ for all diseases $d \in D$, then VFSREP is solvable in $O(\tau(\delta+\delta_{NE})^2)$ time.*

PROOF: Consider some time period $t \in T$. Since the childhood immunization schedule is tight, if dose j for disease $d \in D$ may be administered (i.e., $P_{djt} = 1$), then it must be administered in time period $t \in T$. Let $D_t = \{d \in D: P_{djt} = 1 \text{ for some } j = 1,2,\ldots,n_d\}$ and $V_t = \{v \in V: I_{vd} = 1 \text{ and } d \in D_t\}$ for $t \in T$. Without loss of generality, assume $\Sigma_{v\in V} I_{vd} = 2$ for all diseases $d \in D_t$, otherwise $X_{tv} = 1$ for $v \in V_t$ such that $I_{vd} = 1$. For each disease $d \in D_t$, define a clause $(X_{tv}, X_{tv'})$, where $v, v' \in V_t$ such that $I_{vd} = I_{v'd} = 1$, since $\Sigma_{v\in V} I_{vd} = 2$, and for all diseases $d \in D_t \cap D_{NE}$, define an additional clause $(1 - X_{tv}, 1 - X_{tv'})$. Furthermore, for all diseases $d \in D_{NE} \backslash D_t$, set $X_{tv} = 0$ for all $v \in V_t$ such that $I_{vd} = 1$. Therefore, time period $t \in T$ yields a 2-SAT problem instance with at most υ Boolean variables and at most $\delta+\delta_{NE}$ clauses since every disease (clause) must be "satisfied" by at least one vaccine. The additional clause for each disease $d \in D_t \cap D_{NE}$ ensures that exactly one vaccine is administered for that disease. This transformation from VFSREP to 2-SAT is linear in the number of clauses $m \leq \delta+\delta_{NE}$. Since a 2-SAT problem is solvable in $O(m^2)$ time, then finding the feasible vaccine set that restricts extraimmunization in time period $t \in T$ is solvable in $O((\delta+\delta_{NE})^2)$ time. Applying this result for all time periods $t \in T$ equates to at most τ 2-SAT problem instances, and hence, the overall complexity for this special case is $O(\tau(\delta+\delta_{NE})^2)$. ■

LEMMA 8: *Given an aligned or completely aligned childhood immunization schedule, if $\Sigma_{v\in V} I_{vd} \leq 2$ for all diseases $d \in D$, then VFSREP is solvable in $O(\tau(\delta+\delta_{NE})^2)$ time or $O((\delta+\delta_{NE})^2)$ time, respectively.*

PROOF: Suppose the childhood immunization schedule is completely aligned, and consider some time period $t \in T$ such that $P_{djt} = 1$ for some dose $j = 1,2,\ldots,n_d$ and disease $d \in D$. Since the childhood immunization schedule is completely aligned, then $P_{djt} = 1$ for all diseases $d \in D$. Without loss of generality, assume $\Sigma_{v\in V} I_{vd} = 2$ for all diseases $d \in D$, otherwise $X_{tv} = 1$ for $v \in V$ such that $I_{vd} = 1$. For each disease $d \in D$, define a clause $(X_{tv}, X_{tv'})$, where $v, v' \in V$ such that $I_{vd} = I_{v'd} = 1$, since $\Sigma_{v\in V} I_{vd} = 2$, and for all diseases $d \in D_{NE}$, define an additional clause $(1 - X_{tv}, 1 - X_{tv'})$. Therefore, time period $t \in T$ yields a 2-SAT problem instance with υ Boolean variables and $\delta+\delta_{NE}$ clauses since every disease (clause) must be "satisfied" by at least one vaccine. The additional clause for each disease $d \in D_{NE}$ ensures that exactly one vaccine is administered for that disease. This transformation from VFSREP to 2-SAT is linear to the number of clauses $m \leq \delta+\delta_{NE}$. Since a 2-SAT problem is solvable in $O(m^2)$ time, then finding the feasible vaccine set that restricts extraimmunization in time period $t \in T$ is solvable in $O((\delta+\delta_{NE})^2)$ time. Furthermore, each time period $t \in T$ such that $P_{djt} = 1$ yields an identical 2-SAT problem instance, and hence, requires a single solve. The solution of this 2-SAT problem instance is then applied in some time period such that $P_{djt} = 1$ for all diseases $d \in D$ and doses $j = 1,2,\ldots,n_d$. Therefore, the complexity for this special case is $O((\delta+\delta_{NE})^2)$. Similarly, if the childhood immunization is aligned, then there are at most τ unique 2-SAT problem instances, and hence, the overall complexity for this special case is $O(\tau(\delta+\delta_{NE})^2)$. ■

Lemmas 6, 7, and 8 and Theorem 9 imply VFSREP is solvable in polynomial time if the vaccine set is restricted and if the childhood immunization schedule is tight, aligned, or completely aligned. Moreover, when $\delta = 1$, VFSREP is solvable in polynomial time (i.e., $O(\tau)$ time) since it is a special case of Lemma 6. VFSREP is also solvable in polynomial time when $\delta = 2$ using the dynamic programming algorithm presented in Section 4.3.1, since the subproblem solved at each stage of the dynamic program can be solved as a 2-SAT problem instance, which is solvable in polynomial time (Garey and Johnson 1979).

Table 5 summarizes the complexity results for VFSREP, where $\boldsymbol{n}$ = the dose requirement for each disease $d \in D$ (i.e., $n_d = \boldsymbol{n}$ for all $d \in D$).

Table 5: Summary of Complexity Results for VFSREP

	Time Periods	Diseases—(General Set)	Diseases—(Extraimmunization allowed)	Diseases—(with no Extraimmunization)	Vaccines	# of Doses for each Disease
	$S = T$	$S = D$	$S = D_E$	$S = D_{NE}$	$S = V$	$n = \lvert S \rvert$
$S = \emptyset$	Undefined	Undefined	*NP*-Comp	Polynomial	Infeasible	Polynomial
$\lvert S \rvert = 1$	*NP*-comp	Polynomial	⇓	*NP*-comp	*NP*-comp	*NP*-comp
$\lvert S \rvert = 2$	⇓	Polynomial		⇓	⇓	⇓
$\lvert S \rvert \geq 3$		*NP*-comp				

4.3 Algorithms and Heuristics

This section discusses both exact algorithms and heuristics for VFSREP(O). Section 4.3.1 presents an exact dynamic programming algorithm for VFSREP(O). Section 4.3.2 presents two rounding heuristics (*Rounding* and *MAX Rounding*) for VFSREP(O)-MED. Section 4.3.3 presents a *Greedy* heuristic for VFSREP(O)-MED. Section 4.3.4 presents a *Randomized Rounding* heuristic for VFSREP(O)-MED(A). Lastly, Section 4.3.5 presents a *MAX Rounding* and *Greedy* heuristic for VFSREP(O). Some of the heuristics are shown to be approximation algorithms, which provide an approximation bound on the value of the heuristic solution.

4.3.1 Dynamic Programming Algorithm

In Section 4.1, VFSREP(O) is modeled as an integer programming (IP) problem, and hence, may be solved using several well-known integer optimization techniques (such as branch and bound; see Nemhauser and Wolsey 1999). Another useful exact algorithm is dynamic programming (DP). This section presents and analyzes a DP algorithm for VFSREP(O) similar to the DP presented in section 3.3.1 for VFSLBP(O).

Given the stated set of inputs for VFSREP(O) (i.e., set of time periods T, set of diseases D, D_E, and D_{NE}, set of vaccines V, required number of doses n_d for each $d \in D$, and binary parameters I, P, Q, and R), the DP algorithm solves VFSREP(O) one period at a time beginning at the first time period (i.e., $t = 1$), and steps through each time period in T until $t = \tau$. Therefore, the set T defines the stages of the DP algorithm. In addition to the minimum dose parameter m_{dt}, $d \in D$, $t \in T$, define M_{dt} as the maximum number of doses of a vaccine required for disease $d \in D$ through time period $t \in T$.

Define a state in the DP algorithm as the number of required doses of a vaccine that have been administered for each disease through time period $t \in T$. Formally, a state in time period $t \in T$ is a δ-dimensional vector $\mathbf{S}_t = (S_{t1}, S_{t2}, \ldots, S_{t\delta})$, where S_{td} is the number of required doses of a vaccine that have been administered for disease $d = 1,2,\ldots,\delta$, in time periods $1,2,\ldots,t$. Therefore, the state space in time period $t \in T$ is $\Omega_t = \{\mathbf{S}_t \in \mathbf{Z}^\delta : m_{dt} \leq S_{td} \leq M_{dt}$ for all $d \in D\}$, where $\mathbf{Z}$ denotes the set of all integers. The decision in time period $t \in T$ is which vaccines to administer that immunize against the diseases requiring vaccination in this time period (i.e., the binary decision variables X_{tv}), and is represented by the δ-dimensional binary vector $\mathbf{Y}_t = (Y_{t1}, Y_{t2}, \ldots, Y_{t\delta})$, where $Y_{td} = 1$ implies $X_{tv} = 1$ for some vaccine $v \in V$ that immunizes against disease $d \in D$ (i.e., $I_{vd} = 1$). The decision space in time period $t \in T$ is defined as $\Phi_t = \{\mathbf{Y}_t \in \mathbf{B}^\delta : 0 \leq Y_{td} \leq M_{dt} - m_{d(t-1)}$ for all $d \in D\}$, where $\mathbf{B}$ denotes the binary set $\{0,1\}$. These states and decisions define the DP algorithm system dynamics: $\mathbf{S}_t = \mathbf{S}_{t-1} + \mathbf{Y}_t$. Since $\mathbf{Y}_t \in \Phi_t$ is a binary vector, a state $\mathbf{S}_t \in \Omega_t$ is accessible from state $\mathbf{S}_{t-1} \in \Omega_{t-1}$ only if $\mathbf{S}_t - \mathbf{S}_{t-1}$ is also a binary vector. Furthermore, $\mathbf{Y}_t \in \Phi_t$ being binary eliminates the necessity of the binary decision variables U_{dt}, $d \in D$, $t \in T$, because the vaccines administered in time period $t \in T$ satisfy at most one dose for a particular disease.

Given that $\mathbf{Y}_t = \mathbf{S}_t - \mathbf{S}_{t-1}$, then a transition from state $\mathbf{S}_{t-1} \in \Omega_{t-1}$ to state $\mathbf{S}_t \in \Omega_t$ requires that a dose of vaccine be administered in time period $t \in T$ for each disease in the set $D_t = \{d$

$\in D : Y_{td} = 1\}$. The sets $V_t = \{v \in V: I_{vd} = 1$ and $d \in D_t\}$ (i.e., the set of vaccines that immunize against any disease that requires vaccination in time period $t \in T$) and D_t define a sub-instance of VFSREP(O), termed IP($\mathbf{Y}_t$). To describe IP($\mathbf{Y}_t$), the following definitions are needed. Let

- $D_{Et} = D_E \cap D_t$ and $D_{NEt} = D_{NE} \cap D_t$ for any time period $t \in T$,
- $Z_{dt}^P \in \mathbf{Z}^+ \cup \{0\}$ be the number of extra doses of vaccine administered for disease $d \in D_{NE}$ in time period $t \in T$ such that $Y_{td} = 1$,
- $Z_{dt}^R \in \mathbf{Z}^+ \cup \{0\}$ be the number of extra doses of vaccine administered for disease $d \in D_{NE}$ in time period $t \in T$ such that $Y_{td} = 0$, (i.e., for disease $d \in D_{NE} \backslash D_{NEt}$).

The specific sub-instance for VFSREP(O) for time period $t \in T$ and decision $\mathbf{Y}_t \in \Phi_t$ is given by

IP($\mathbf{Y}_t$)

$$\text{Minimize} \quad \sum_{d \in D_{NEt}} \rho_d Z_{dt}^P + \sum_{d \in D_{NE} \backslash D_{NEt}} \gamma_d Z_{dt}^R$$

Subject to

$$\sum_{v \in V_t} X_{tv} I_{vd} \geq 1 \qquad \text{for all } d \in D_{Et},$$

$$\sum_{v \in V_t} X_{tv} I_{vd} - Z_{dt}^P = 1 \qquad \text{for all } d \in D_{NEt},$$

$$\sum_{v \in V_t} X_{tv} I_{vd} - Z_{dt}^R = 0 \qquad \text{for all } d \in D_{NE} \backslash D_{NEt},$$

$$X_{tv} \in \{0,1\} \qquad \text{for all } v \in V_t,$$

$$Z_{dt}^P, Z_{dt}^R \in \mathbf{Z}^+ \cup \{0\} \qquad \text{for all } d \in D_{NE}.$$

To characterize the cost of decision $\mathbf{Y}_t \in \Phi_t$, which is the cost of transitioning from state $\mathbf{S}_{t-1} \in \Omega_{t-1}$ in time period $(t-1) \in T$ to state $\mathbf{S}_t \in \Omega_t$ in time period $t \in T$, define the one-period cost function $C_t(\mathbf{S}_{t-1}, \mathbf{Y}_t)$ as the amount of extraimmunization in time period $t \in T$ given state $\mathbf{S}_{t-1} \in \Omega_{t-1}$ and decision $\mathbf{Y}_t \in \Phi_t$. Note, however, that this one-period cost in time period $t \in T$ depends only on decision $\mathbf{Y}_t \in \Phi_t$, and hence, the optimal value of IP($\mathbf{Y}_t$) = $C_t(\mathbf{S}_{t-1}, \mathbf{Y}_t) = C_t(\mathbf{Y}_t)$. Therefore, the optimal one-period value over all possible decisions in time period $t \in T$ is given by $\min_{\mathbf{Y}_t \in \Phi_t} C_t(\mathbf{Y}_t)$.

Define $Z_t(\mathbf{S}_t)$ as the minimum weighted (as defined by ρ_d and γ_d for disease $d \in D_{NE}$) amount of extraimmunization of a vaccine formulary that immunizes against all diseases through time period $t \in T$ subject to the number of required doses at the end of time period t

$\in T$ being equal to $\mathbf{S}_t \in \Omega_t$. Therefore, the DP optimality equation is given by the recurrence relation

$$Z_t(\mathbf{S}_t) = \min_{\mathbf{Y}_t \in \Phi_t, \mathbf{S}_{t-1} \in \Omega_{t-1} : \mathbf{S}_t = \mathbf{S}_{t-1} + \mathbf{Y}_t} \{C_t(\mathbf{Y}_t) + Z_{t\text{-}1}(\mathbf{S}_{t\text{-}1})\}.$$

Furthermore, the minimum weighted amount of extraimmunization that satisfies a given childhood immunization schedule is given by

$$z^* = \min_{\mathbf{S}_\tau \in \Omega_\tau} Z_t(\mathbf{S}_t),$$

where Ω_τ is the state space for the final time period $\tau \in T$. The DP algorithm for VFSREP(O) is now formally given.

Dynamic Programming Algorithm for VFSREP(O)

Step 1. Initialize:
  a. Initial state, $\mathbf{S}_0 \leftarrow \mathbf{0}$ (the δ-dimensional zero vector)
  b. Initial extraimmunization contribution, $Z_0(\mathbf{S}_0) \leftarrow 0$
  c. Set m_{d0}, $M_{d0} \leftarrow 0$ for all $d \in D$
  d. Initial stage, $t \leftarrow 1$

Step 2. Compute

$$Z_t(\mathbf{S}_t) = \min_{\mathbf{Y}_t \in \Phi_t, \mathbf{S}_{t-1} \in \Omega_{t-1} : \mathbf{S}_t = \mathbf{S}_{t-1} + \mathbf{Y}_t} \{C_t(\mathbf{Y}_t) + Z_{t\text{-}1}(\mathbf{S}_{t\text{-}1})\}$$

for each state $\mathbf{S}_t \in \Omega_t$.

Step 3. If $t < \tau$, then $t \leftarrow t + 1$ and return to Step 2. Else, stop and return $z^* = \min_{\mathbf{S}_\tau \in \Omega_\tau} Z_t(\mathbf{S}_t)$.

Example 9

This demonstrates the DP algorithm for VFSREP(O) for the childhood immunization schedule depicted in Figure 5.

Recall $D_{NE} = \{1\}$, $V = \{1 = \{1,2,3\}\}$, and the minimum dose vectors for each disease $d \in D$ are $m_1 = (0,0,1,2,3,3,3,4)$, $m_2 = (0,0,1,2,2,2,2,3)$, and $m_3 = (0,0,0,0,0,0,0,1)$, where m_{dt} is the t^{th} component, $t = 1,2,\ldots,8$, of vector m_d for disease $d = 1,2,3$. Likewise, the maximum dose vectors for each disease $d \in D$ are $M_1 = (0,0,1,2,3,3,4,4)$, $M_2 = (0,0,1,2,3,3,3,3)$, and $M_3 = (0,0,0,0,0,1,1,1)$, where M_{dt} is the t^{th} component, $t = 1,2,\ldots,8$, of vector M_d for disease $d = 1,2,3$. These parameters yield the following state and decision spaces:

State Space for each Time Period							
Ω_1	Ω_2	Ω_3	Ω_4	Ω_5	Ω_6	Ω_7	Ω_8
{(0,0,0)}	{(0,0,0)}	{(1,1,0)}	{(2,2,0)}	{(3,2,0), (3,3,0)}	{(3,2,0), (3,2,1),	{(3,2,0), (3,2,1),	{(4,3,1)}

					(3,3,0), (3,3,1)}	(3,3,0), (3,3,1)} (4,2,0), (4,2,1), (4,3,0), (4,3,1)}	
Decision Space for each Time Period							
Φ_1	Φ_2	Φ_3	Φ_4	Φ_5	Φ_6	Φ_7	Φ_8
{(0,0,0)}	{(0,0,0)}	{(1,1,0)}	{(1,1,0)}	{(1,0,0), (1,1,0)}	{(0,0,0), (0,0,1), (0,1,0), (0,1,1)}	{(0,0,0), (0,0,1), (0,1,0), (0,1,1), (1,0,0), (1,0,1), (1,1,0), (1,1,1)}	{(0,0,0), (0,0,1), (0,1,0), (0,1,1), (1,0,0), (1,0,1), (1,1,0), (1,1,1)}.

Note that for the given disease and vaccine set,

$$C_t(Y_t) = \begin{cases} 1 & \text{when } Y_t = (0, a, b) \text{ where } a, b \in \{0,1\}, a = b \neq 0 \\ 0 & \text{when } Y_t = (0,0,0) \text{ or } Y_t = (1, a, b) \text{ where } a, b \in \{0,1\} \end{cases} \quad \text{for all } t \in T.$$

Applying the DP algorithm, where $\Omega_0 = \{\mathbf{S}_0\} = \{(0,0,0)\}$ and $Z_0(\mathbf{S}_0) = 0$, implies for time period (stage):

$t = 1$:

$Z_1((0,0,0) = C_1((0,0,0)) + Z_0((0,0,0)) = 0 + 0 = 0,$

$t = 2$:

$Z_2((0,0,0)) = C_2((0,0,0)) + Z_1((0,0,0)) = 0 + 0 = 0,$

$t = 3$:

$Z_3((1,1,0)) = C_3((1,1,0) + Z_2((0,0,0)) = 0 + 0 = 0,$

$t = 4$:

$Z_4((2,2,0)) = C_4((1,1,0)) + Z_3((1,1,0)) = 0 + 0 = 0,$

$t = 5$:

$Z_5((3,2,0)) = C_5((1,0,0)) + Z_4((2,2,0)) = 0 + 0 = 0,$

$Z_5((3,3,0)) = C_5((1,1,0)) + Z_4((2,2,0)) = 0 + 0 = 0,$

$t = 6$:

$Z_6((3,2,0)) = C_6((0,0,0)) + Z_5((3,2,0)) = 0 + 0 = 0$

$Z_6((3,3,0)) = \min\{C_6((0,0,0)) + Z_5((3,3,0)); C_6((0,1,0)) + Z_5((3,2,0))\} = \min\{0+0, 1+0\} = 0,$

$Z_6((3,2,1)) = C_6((0,0,1)) + Z_5((3,2,0)) = 1 + 0 = 1,$

$Z_6((3,3,1)) = \min\{C_6((0,0,1) + Z_5((3,3,0)); C_6((0,1,1) + Z_5((3,2,0))\} = \min\{1+0,1+0\} = 1,$

$t = 7$:

$Z_7((3,2,0)) = C_7((0,0,0)) + Z_6((3,2,0)) = 0 + 0 = 0,$

$Z_7((3,2,1)) = \min\{C_7((0,0,1)) + Z_6((3,2,0)); C_7((0,0,0))+ Z_6((3,2,1))\}= \min\{1+0,0+1\}= 1,$

$Z_7((3,3,0)) = \min\{C_7((0,0,0)) + Z_6((3,3,0));C_7((0,1,0))+ Z_6((3,2,0))\} = \min\{0+0,1+0\}= 0,$

$Z_7((3,3,1)) = \min\{C_7((0,0,0)) + Z_6((3,3,1));C_7((0,1,1)) + Z_6((3,2,0)); C_7((0,1,0)) +$
$Z_6((3,2,1)); C_7((0,0,1) + Z_6((3,3,0))\} = \min\{0 + 1,1 + 0,1 + 1,1 + 0\} = 1,$

$Z_7((4,2,0)) = C_7((1,0,0)) + Z_6((3,2,0)) = 0 + 0 = 0,$

$Z_7((4,2,1)) = \min\{C_7((1,0,1)) + Z_6((3,2,0));C_7((1,0,0)) + Z_6((3,2,1)) = \min\{0+0,0+1\} = 0,$

$Z_7((4,3,0)) = \min\{C_7((1,1,0)) + Z_6((3,2,0));C_7((1,0,0))+ Z_6((3,3,0))\} = \min\{0+0,0+0\}= 0,$

$Z_7((4,3,1)) = \min\{C_7((1,1,1) + Z_6((3,2,0));C_7((1,1,0) + Z_6((3,2,1));C_7((1,0,1) +$
$Z_6((3,3,0)); {}_7((1,0,0) + Z_6((3,3,1))\} = \min\{0 + 0,0 + 1,0 + 0,0 + 1\} = 0,$

$t = 8$:

$Z_8((4,3,1)) = \min\{C_8((0,0,1) + Z_7((4,3,0));C_8((1,0,1) + Z_7((3,3,0));C_7((1,1,1) +$
$Z_7((3,2,0));C_8((0,1,1) + Z_7((4,2,0));C_8((1,1,0) + Z_7((3,2,1));C_8((1,0,0) +$
$Z_7((3,3,1));C_8((0,1,0) + Z_7((4,2,1));C_8((0,0,0) + Z_7((4,3,1))\}$
$= \min\{1 + 0,0 + 0,0 + 0,1 + 0,0 + 1,0 + 1,1 + 0,0 + 0\} = 0.$

Therefore, the minimum amount of extraimmunization for this vaccine formulary is 0 (i.e., there exists a feasible immunization schedule that restricts extraimmunization). Furthermore, one possible vaccination schedule (highlighted above) that restricts extraimmunization is to administer vaccine $v = 1$ in time periods 3, 4, 5, and 7. □

To determine the complexity of this DP algorithm, suppose that the IP($\mathbf{Y}_t$) problem instance with δ diseases and υ vaccines can be solved in $O(\mathbf{T}_{\text{IP}})$ time. Furthermore, define $\mathbf{S}_{Max}$ to be the maximum number of possible states within any time period $t \in T$. Each time period requires $O((\mathbf{S}_{Max})^2 \cdot \mathbf{T}_{\text{IP}})$ time, and hence, with τ time periods, the DP algorithm for VFSREP(O) executes in $O(\tau(\mathbf{S}_{Max})^2 \cdot \mathbf{T}_{\text{IP}})$ time. The fact that IP($\mathbf{Y}_t$) for $\tau = 1$ with δ diseases and υ vaccines is NP-hard follows from Theorem 7, and hence, a polynomial (or even pseudo-polynomial) algorithm is unlikely to exist, unless $P = NP$. The DP algorithm's worst case complexity may be improved, however, since each IP($\mathbf{Y}_t$) instance depends only on the decision vector $\mathbf{Y}_t \in \Phi_t$. Therefore, IP($\mathbf{Y}_t$) for decision $\mathbf{Y}_t \in \Phi_t$ only needs to be solved once.

It can be shown that the complexity of solving for all possible decisions is $O(\upsilon\delta 2^{\delta})$. This means that for each time period $t \in T$, the complexity of Step 2 becomes $O(\delta(\mathbf{S}_{Max})^2)$, and hence, the DP algorithm has a $O(\tau\delta(\mathbf{S}_{Max})^2 + \upsilon\delta 2^{\delta})$ worst case time complexity, which is an improvement over $O(\tau(\mathbf{S}_{Max})^2 \cdot \mathbf{T}_{\text{IP}})$ when $\mathbf{S}_{Max}$ is large. To exploit this added efficiency, the implementation of the DP algorithm used for the computational analysis reported in Section 4.4 employs a 'branch and remember' recursive algorithm to find the optimal value for each IP($\mathbf{Y}_t$) instance. Therefore, IP($\mathbf{Y}_t$) need only be computed once using the recursive algorithm *Single-Period-IP*. This recursive algorithm assumes $\rho_d = \gamma_d$ for all diseases $d \in D_{NE}$. Initially, the given set of diseases for $\mathbf{Y}_t$ is D_t, and hence, $D' = D_t$.

Single-Period-IP(D')
- If $D' = \emptyset$, return 0 as the solution value
- If IP($\mathbf{Y}_t$) for $D' = \{d \in D : Y_{td} = 1\}$ has been solved previously, return its optimal value
- Select a disease $d \in D'$ that requires immunization
- Let $V' = \{v \in V: I_{vd} = 1\}$ (set of vaccines $v \in V$ that immunize against disease $d \in D'$)
- Set $LowestPenalty = +\infty$
- For each vaccine $v \in V'$
  - Let $D^* = D' \setminus \{d \in D': I_{vd} = 1\}$
  - $Penalty$ = *Single-Period-IP*(D^*) (find the optimal penalty for the set of diseases D^*)
  - Let $D_{NEv} = \{d \in \overline{D'} : d \in D_{NE}, I_{vd} = 1\}$
  - Set $Penalty_v = 0$
  - For each disease $d \in D_{NEv}$
    - $Penalty_v = Penalty_v + \gamma_d$
  - If $Penalty + Penalty_v < LowestPenalty$
    - $LowestPenalty = Penalty + Penalty_v$
  - Store $LowestPenalty$ for D' (save the optimal solution for the set of diseases D')
- Return $LowestPenalty$

Despite its exponential worst case complexity run time, the DP algorithm for VFSREP(O) offers several advantages as described in Section 3.3.1 for VFSLBP(O), and is efficient in practice (see Section 4.4) with the 2006 Recommended Childhood Immunization Schedule, since this schedule yields a reasonable state space, and the IP($\mathbf{Y}_t$) instances in each time period $t \in T$ are small (and, in many cases, are solvable in polynomial time).

4.3.2 Rounding Heuristics

The worst case complexity for the DP algorithm motivates the need for heuristics. This section presents the *Rounding* and *MAX Rounding* heuristics for VFSREP(O)-MED. VFSREP(O)-MED is first considered due to its simpler structure and its relation to the 2006

Recommended Childhood Immunization Schedule (all diseases have mutually exclusive doses). Both *Rounding* and *MAX Rounding* are shown to be approximation algorithms, which, by definition, execute in polynomial time and provide an approximation bound on the value of the heuristic solution (Hochbaum 1997).

The *Rounding* and *MAX Rounding* heuristics use the solution from a linear program (LP) to construct a feasible binary solution. This technique has been applied to several other well-known discrete optimization problems such as the Set-Covering problem (Hochbaum 1997). Relaxing the binary and integer constraints for the decision variables in VFSREP(O)-MED yields the LP relaxation

$$\text{Minimize} \quad \sum_{d \in D_{NE}} \left(\rho_d \left(\sum_{j=1}^{n_d} Z_{dj}^{P} \right) + \gamma_d Z_d^{R} \right)$$

Subject to

$$\sum_{t \in T} \sum_{v \in V} P_{djt} X_{tv} I_{vd} \geq 1 \qquad \text{for all } d \in D_E,\ j = 1,2,\ldots,n_d,$$

$$\sum_{t \in T} \sum_{v \in V} P_{djt} X_{tv} I_{vd} - Z_{dj}^{P} = 1 \qquad \text{for all } d \in D_{NE},\ j = 1,2,\ldots,n_d,$$

$$\sum_{t \in T} \sum_{v \in V} R_{dt} X_{tv} I_{vd} - Z_d^{R} = 0 \qquad \text{for all } d \in D_{NE},$$

$$0 \leq X_{tv} \leq 1 \qquad \text{for all } t \in T,\ v \in V,$$

$$Z_{dj}^{P} \geq 0 \qquad \text{for all } d \in D_{NE},\ j = 1,2,\ldots,n_d,$$

$$Z_d^{R} \geq 0 \qquad \text{for all } d \in D_{NE}.$$

Denote the optimal objective function values of VFSREP(O)-MED and its LP relaxation as z_{IP} and z_{LP}, respectively, where $z_{LP} \leq z_{IP}$ (since the feasible region of VFSREP(O)-MED is contained in the feasible region of its LP relaxation). Let X_{LP}^{*} denote the optimal decision vector for the LP relaxation and $X_{LP\,tv}^{*}$, $t \in T$, $v \in V$, Z_{dj}^{P*}, $d \in D_{NE}$, j=1,2,…,n_d, and Z_d^{R*}, $d \in D_{NE}$, denote the optimal values for the decision variables in the LP relaxation. Let $\alpha_d \equiv (\Sigma_{v \in V} I_{vd})(\max_{j=1,2,\ldots,n_d} \Sigma_{t \in T} P_{djt})$ for all diseases $d \in D$, the maximum number of non-zero columns in any row of the constraint matrix for VFSREP(O)-MED corresponding to disease $d \in D$, and $\alpha \equiv \max_{d \in D} \alpha_d$. The *Rounding* heuristic rounds each fractional decision variable $X_{LP\,tv}^{*}$, $t \in T$, $v \in V$, which is greater than the threshold value $1/\alpha$, and then computes the weighted amount of extraimmunization using these rounded variables. The *Rounding* heuristic is now formally given.

Rounding Heuristic for VFSREP(O)-MED

Step 1. Solve the LP relaxation of VFSREP(O)-MED
Step 2. $X_{tv} \leftarrow 0$ for all $t \in T$ and $v \in V$
Step 3. For all $t \in T$ and $v \in V$
  a. If $X^*_{LP_{tv}} \geq 1/\alpha$, then $X_{tv} \leftarrow 1$
Step 4. For all $d \in D_{NE}$
  a. For all $j = 1,2,\ldots,n_d$
    i. $Z^P_{dj} \leftarrow \Sigma_{t\in T}\Sigma_{v\in V}\, P_{djt}\, X_{tv}\, I_{vd} - 1$
  b. $Z^R_d \leftarrow \Sigma_{t\in T}\Sigma_{v\in V}\, R_{dt}\, X_{tv}\, I_{vd}$
Step 5. Compute and return $\sum_{d\in D_{NE}} \left(\rho_d \left(\sum_{j=1}^{n_d} Z^P_{dj}\right) + \gamma_d Z^R_d\right)$.

Example 10

Consider the childhood immunization schedule displayed in Figure 3 together with the vaccine set $V = \{1 = \{1\}, 2 = \{2\}, 3 = \{2,3\}\}$, disease set $D_{NE} = \{1,2\}$, and weights $\rho_d = \gamma_d = 1$ for diseases $d = 1,2$.

In this example, $z_{LP} = 1$, since a dose of vaccine $v = 3$ must be administered in time period $t = 2$ or 3 to satisfy the dose requirement for disease $d = 3$, and hence, $Z^R_2 = 1$. Therefore, a feasible decision vector for $X^*_{LP} = (X^*_{LP_{11}}, X^*_{LP_{12}}, X^*_{LP_{13}}, X^*_{LP_{21}}, X^*_{LP_{22}}, X^*_{LP_{23}}, X^*_{LP_{31}}, X^*_{LP_{32}}, X^*_{LP_{33}}, X^*_{LP_{41}}, X^*_{LP_{42}}, X^*_{LP_{43}})$ that yields this optimal value is $X^*_{LP} = (1, \frac{5}{6}, \frac{1}{6}, 0, 0, \frac{4}{5}, \frac{3}{4}, 0, \frac{1}{5}, \frac{1}{4}, \frac{2}{3}, \frac{1}{3})$. Furthermore, $\alpha = 4$, and hence, the *Rounding* heuristic rounds all binary variables $\geq$ ¼ yielding the binary assignment (1,1,0,0,0,1,1,0,0,1,1,1), which returns an objective function value of 3, since $Z^P_{12} = 1$, $Z^P_{22} = 1$, and $Z^R_2 = 1$. □

Lemma 9 establishes the feasibility of the solution returned by the *Rounding* heuristic.

LEMMA 9: *The Rounding heuristic for VFSREP(O)-MED returns a feasible binary solution X,* (*i.e., a decision vector that satisfies the childhood immunization schedule*).

PROOF: By way of contradiction, suppose the *Rounding* heuristic does not produce a feasible solution that satisfies the childhood immunization schedule. Then there exists some disease $d \in D$ whose j^{th} dose is not administered during some time period $t \in T$ such that $P_{djt} = 1$ or, for some disease $d \in D_{NE}$, $\Sigma_{t\in T}\Sigma_{v\in V}\, P_{djt}\, X_{tv}\, I_{vd} - Z^P_{dj} \neq 1$ for some dose $j = 1,2,\ldots,n_d$ or $\Sigma_{t\in T}\Sigma_{v\in V}\, R_{dt}\, X_{tv}\, I_{vd} - Z^R_d \neq 0$. However, Step 4 of the algorithm ensures for all diseases $d \in$

D_{NE} that $\Sigma_{t\in T}\Sigma_{v\in V} P_{djt} X_{tv} I_{vd} - Z_{dj}^{P} = 1$ for dose $j = 1,2,\ldots,n_d$ and $\Sigma_{t\in T}\Sigma_{v\in V} R_{dt} X_{tv} I_{vd} - Z_{d}^{R} = 0$. Therefore, for *Rounding* heuristic to yield an infeasible solution there must exists some disease $d \in D$ whose j^{th} dose is not administered during some time period $t \in T$ such that $P_{djt} = 1$. This implies that $X^*_{LP_{tv}} < 1/\alpha$ for all decision variables in the constraint corresponding to disease $d \in D$, dose j. However, by definition of α, there are at most α decision variables in this constraint. Therefore, for disease $d \in D$, dose j

$$\sum_{t\in T}\sum_{v\in V} P_{djt} X^*_{tv} I_{vd} < \alpha(1/\alpha) < 1,$$

which violates the LP relaxation constraint for disease $d \in D$, dose j, but this contradicts the feasibility of X^*_{LP}. ■

Given that linear programming is solvable in polynomial time, it then follows that the *Rounding* heuristic executes in polynomial time. Theorem 10 shows that the value of the binary solution returned by the *Rounding* heuristic is guaranteed to be no worse than $\alpha \cdot z_{IP}$.

THEOREM 10: *The Rounding heuristic is an α-approximation algorithm for VFSREP(O)–MED.*

PROOF: Clearly, the *Rounding* heuristic executes in polynomial time since LP executes in polynomial time.

It remains to show that $\sum_{d\in D_{NE}} (\rho_d \sum_{j=1}^{n_d} Z_{dj}^{P} + \gamma_d Z_d^R) \le \alpha \cdot z_{IP}$. By step 4 of the algorithm,

$$\sum_{d\in D_{NE}} (\rho_d \sum_{j=1}^{n_d} Z_{dj}^{P} + \gamma_d Z_d^R) = \sum_{d\in D_{NE}} (\rho_d \sum_{j=1}^{n_d} (\sum_{t\in T}\sum_{v\in V} P_{djt} X_{tv} I_{vd} - 1) + \gamma_d \sum_{t\in T}\sum_{v\in V} R_{dt} X_{tv} I_{vd})$$

$$\le \sum_{d\in D_{NE}} \sum_{t\in T}\sum_{v\in V} (\rho_d \sum_{j=1}^{n_d} (P_{djt} (X^*_{LP_{tv}} \alpha) I_{vd} - 1) + \gamma_d R_{dt} (X^*_{LP_{tv}} \alpha) I_{vd})$$

(since $X_{tv} = 1$ only if $X^*_{LP_{tv}} \alpha \geq 1$)

$$= \alpha \sum_{d\in D_{NE}} \sum_{t\in T}\sum_{v\in V} (\rho_d \sum_{j=1}^{n_d} (P_{djt} X^*_{LP_{tv}} I_{vd} - 1) + \gamma_d R_{dt} X^*_{LP_{tv}} I_{vd})$$

$$= \alpha \sum_{d\in D_{NE}} (\rho_d \sum_{j=1}^{n_d} (\sum_{t\in T}\sum_{v\in V} P_{djt} X^*_{LP_{tv}} I_{vd} - 1) + \gamma_d \Sigma_{t\in T}\Sigma_{v\in V} R_{dt} X^*_{LP_{tv}} I_{vd})$$

$$= \alpha \sum_{d\in D_{NE}} (\rho_d \sum_{j=1}^{n_d} Z_{dj}^{P*} + \gamma_d Z_d^{R*})$$

$= \alpha \cdot z_{LP}$

$\leq \alpha \cdot z_{IP}$ (since $z_{LP} \leq z_{IP}$). ■

Theorem 10 implies two immediate corollaries for special cases of VFSREP(O)-MED. Corollary 7 considers a tight childhood immunization schedule such that there are at most two vaccines that immunize against each disease $d \in D$ (i.e., $\Sigma_{v\in V} I_{vd} \leq 2$ for all diseases $d \in D$), and Corollary 8 gives an upper bound on α for a tight childhood immunization schedule and for an arbitrary childhood immunization schedule.

COROLLARY 7: *Given a tight childhood immunization schedule, if $\Sigma_{v\in V} I_{vd} \leq 2$ for all diseases $d \in D$, then the Rounding heuristic is a 2-approximation algorithm for VFSREP(O)-MED.*

PROOF: A tight childhood immunization schedule implies $\max_{j=1,2,\ldots,n_d} \Sigma_{t\in T} P_{djt} = 1$ for all diseases $d \in D$. Moreover, $\Sigma_{v\in V} I_{vd} \leq 2$ for all diseases $d \in D$, and hence, $\alpha_d = (\Sigma_{v\in V} I_{vd})(\max_{j=1,2,\ldots,n_d} \Sigma_{t\in T} P_{djt}) \leq 2$ for all diseases $d \in D$. Therefore, $\alpha = \max_{d\in D} \alpha_d \leq 2$. ■

COROLLARY 8: *i*) *Given a tight childhood immunization schedule, $\alpha \leq \upsilon$ for the Rounding heuristic.* *ii*) *Given an arbitrary childhood immunization schedule, $\alpha \leq \upsilon\cdot\tau$ for the Rounding heuristic.*

PROOF: Given a tight childhood immunization schedule, $\max_{j=1,2,\ldots,n_d} \Sigma_{t\in T} P_{djt} = 1$ for all diseases $d \in D$, which implies $\alpha = \max_{d\in D} (\Sigma_{v\in V} I_{vd})(\max_{j=1,2,\ldots,n_d} \Sigma_{t\in T} P_{djt}) \leq \upsilon$. Moreover, for an arbitrary childhood immunization schedule, $\alpha = \max_{d\in D} (\Sigma_{v\in V} I_{vd})(\max_{j=1,2,\ldots,n_d} \Sigma_{t\in T} P_{djt}) \leq \upsilon\cdot\tau$. ■

If X^*_{LP} contains several fractional variables, then the *Rounding* heuristic tends to round too many variables to one, thereby yielding a significant amount of extraimmunization. Instead of rounding all variables greater than or equal to the $1/\alpha$ threshold, it seems reasonable to round only a few variables with large fractional values (i.e., variables closest to one), since these variables are more likely to equal one in the optimal solution. The *MAX Rounding* heuristic limits the number of rounded variables by selecting the variables with large fractional values.

To present the *MAX Rounding* heuristic, recall the notation $\mathbf{D} = \{(d, j): d \in D, j = 1,2,\ldots,n_d\}$ to be the set of all diseases ordered by dose, where $|\mathbf{D}| = \Sigma_{d=1}^{\delta} n_d$, and, for all time periods $t \in T$ and vaccines $v \in V$, $C_{tv} = \{(d, j) \in \mathbf{D}: I_{vd} = 1 \text{ and } P_{djt} = 1\}$, which specifies the diseases and dose that vaccine $v \in V$ immunizes against in time period $t \in T$. Therefore, $C_{tv} \subseteq \mathbf{D}$ for all time periods $t \in T$ and vaccines $v \in V$. Furthermore, in the case when all diseases $d \in D$ have mutually exclusive doses, at most one $(d, j) \in \mathbf{D}$ for all diseases $d \in D$ is contained in any set C_{tv}, since for a given disease $d \in D$ and time period $t \in T$, $P_{djt} = 1$ for at most one dose $j = 1,2,\ldots,n_d$, and hence, each set C_{tv} does not contain multiple doses for any disease $d \in D$. Lastly, define $f_{tv} = X^*_{LP_{tv}}$ for all time periods $t \in T$ and vaccines $v \in V$, which specifies the value of vaccine $v \in V$ in time period $t \in T$. Therefore, the *MAX Rounding* heuristic limits the number of rounded variables by greedily selecting (at each iteration) the most valuable available vaccine $v \in V$ that immunizes against the most disease doses (not yet covered) in time period $t \in T$ (i.e., rounds the variable $X^*_{LP_{tv}}$ that maximizes $f_{tv} \cdot |C_{tv}|$) until every disease dose $(d, j) \in \mathbf{D}$ is covered by some vaccine $v \in V$ in time period $t \in T$. The *MAX Rounding* heuristic is now formally given.

MAX Rounding Heuristic for VFSREP(O)-MED

Step 1. Initialize:

a. Solve the LP relaxation of VFSREP(O)-MED
b. $f_{tv} \leftarrow X^*_{LP_{tv}}$ for all $t \in T$, $v \in V$ such that $X^*_{LP_{tv}} \geq 1/\alpha$
c. $X_{tv} \leftarrow 0$ for all $t \in T$ and $v \in V$
d. $\hat{C}_{tv} \leftarrow C_{tv}$ for all $t \in T$ and $v \in V$

Step 2. While $\mathbf{C} = \bigcup_{\{tv: X_{tv}=1\}} C_{tv} \neq \mathbf{D}$ do

a. $(t', v') \leftarrow \arg\max_{t \in T, v \in V} f_{tv} \cdot |\hat{C}_{tv}|$ (select the non-empty set $\hat{C}_{tv}$ with the largest fractional value times the number of disease doses covered by vaccine $v \in V$ in time period $t \in T$)
b. $X_{t'v'} \leftarrow 1$ (administer vaccine $v' \in V$ in time period $t' \in T$)
c. $\hat{C}_{tv} \leftarrow \hat{C}_{tv} \setminus \hat{C}_{t'v'}$ for all $t \in T$ and $v \in V$ (remove all the disease doses covered by vaccine $v' \in V$ in time period $t' \in T$ from all remaining sets)

Step 3. For all $d \in D_{NE}$

a. For all $j = 1,2,\ldots,n_d$
   i. $Z^P_{dj} \leftarrow \Sigma_{t \in T} \Sigma_{v \in V} P_{djt} X_{tv} I_{vd} - 1$
b. $Z^R_d \leftarrow \Sigma_{t \in T} \Sigma_{v \in V} R_{dt} X_{tv} I_{vd}$

Step 4. Compute and return $\sum_{d \in D_{NE}} \left(\rho_d \left(\sum_{j=1}^{n_d} Z^P_{dj} \right) + \gamma_d Z^R_d \right)$.

Example 11

Consider Example 4 for the *MAX Rounding* heuristic for VFSLBP(O), which is for childhood immunization schedule displayed in Figure 3 together with the sets and weights given in Example 10. The *MAX Rounding* heuristic for Steps 1 and 2 proceeds as given in Example 4, and returns the binary assignment (1,1,0,0,0,1,1,0,0,0,1,0). Step 3 then calculates the values $Z_{11}^{P} = 0$, $Z_{12}^{P} = 0$, $Z_{21}^{P} = 0$, $Z_{22}^{P} = 0$, and $Z_{2}^{R} = 1$, and Step 4 computes and returns the objective function value of 1, which is the optimal amount of extraimmunization. □

The *MAX Rounding* heuristic executes in $O(\mathbf{T}_{\mathrm{LP}} + |\mathbf{D}|\tau\upsilon)$ time, where $\mathbf{T}_{\mathrm{LP}}$ is the time required to solve the LP relaxation of VFSREP(O)-MED. Furthermore, the *MAX Rounding* heuristic returns a feasible solution, since every iteration of the while loop (i.e., Step 2) administers a vaccine that satisfies at least one dose requirement for some disease $d \in D$ (i.e., every iteration covers at least one $(d, j) \in \mathbf{D}$). Moreover, Step 1.b. ensures that the solution returned by the *MAX Rounding* heuristic can be no worse than the solution returned by the *Rounding* heuristic, and hence, the *MAX Rounding* heuristic is also an α-approximation algorithm for VFSREP(O)-MED.

4.3.3 Greedy Heuristic

This section presents the *Greedy* heuristic for VFSREP(O)-MED. The *Greedy* heuristic iteratively selects the vaccine that incurs the smallest penalty for extraimmunization and immunizes against the most disease doses. Recall, $\mathbf{D} = \{(d, j) : d \in D, j = 1,2,\ldots,n_d\}$, $C_{tv} = \{(d, j) \in \mathbf{D}: I_{vd} = 1 \text{ and } P_{djt} = 1\}$ for all time periods $t \in T$ and vaccines $v \in V$, and define the extraimmunization penalty for vaccine $v \in V$ in time period $t \in T$ as $W_{tv} = \sum_{\{d \in D: I_{vd}=1\}} w_{dt}$, where

$$w_{dt} = \begin{cases} \rho_d & \text{if } d \in D_{NE}, (d, j) \in C_{tv} \text{ for some } j = 1,2,\ldots,n_d \text{, and } (d, j) \in \mathbf{C} = \bigcup_{\{tv: X_{tv}=1\}} C_{tv} \\ \gamma_d & \text{if } d \in D_{NE}, (d, j) \notin C_{tv} \text{ for some } j = 1,2,\ldots,n_d \\ 0 & \text{otherwise,} \end{cases}$$

since the penalty for extraimmunization is ρ_d if dose requirement $j = 1,2,\ldots,n_d$ for disease $d \in D_{NE}$ is satisfied by some vaccine in an earlier iteration, γ_d if vaccine $v \in V$ immunizes against disease $d \in D_{NE}$ but does not satisfy some dose requirement in time period $t \in T$, or zero for all diseases $d \in D_{NE}$ such that vaccine $v \in V$ does not provide immunization, (i.e., $I_{vd} = 0$) and for all diseases $d \in D_E$. The *Greedy* heuristic is now formally given.

Greedy Heuristic for VFSREP(O)-MED

Step 1. Initialize:

a. $X_{tv} \leftarrow 0$ for all $t \in T$ and $v \in V$

b. $\hat{C}_{tv} \leftarrow C_{tv}$ for all $t \in T$ and $v \in V$

Step 2. While $\mathbf{C} = \bigcup_{\{tv:X_{tv}=1\}} C_{tv} \neq \mathbf{D}$ do

a. Compute W_{tv} for all $t \in T$ and $v \in V$ (compute extraimmunization penalty for vaccine $v \in V$ in time period $t \in T$)

b. $(t', v') \leftarrow \arg\min_{t\in T, v\in V} W_{tv} / |\hat{C}_{tv}|$ (select the non-empty set $\hat{C}_{tv}$ with the smallest penalty per disease doses covered by vaccine $v \in V$ in time period $t \in T$. Break ties by selecting vaccine $v \in V$ that immunizes against the most diseases in time period $t \in T$.)

c. $X_{t'v'} \leftarrow 1$ (administer vaccine $v' \in V$ in time period $t' \in T$)

d. $\hat{C}_{tv} \leftarrow \hat{C}_{tv} \setminus \hat{C}_{t'v'}$ for all $t \in T$ and $v \in V$ (remove all the disease doses covered by vaccine $v' \in V$ in time period $t' \in T$ from all remaining sets)

Step 3. For all $d \in D_{NE}$

a. For all $j = 1,2,\ldots,n_d$

ii. $Z_{dj}^{P} \leftarrow \Sigma_{t\in T}\Sigma_{v\in V} P_{djt} X_{tv} I_{vd} - 1$

b. $Z_{d}^{R} \leftarrow \Sigma_{t\in T}\Sigma_{v\in V} R_{dt} X_{tv} I_{vd}$

Step 4. Compute and return $\sum_{d\in D_{NE}} \left(\rho_d \left(\sum_{j=1}^{n_d} Z_{dj}^{P}\right) + \gamma_d Z_{d}^{R}\right)$.

Example 12

Consider the childhood immunization schedule displayed in Figure 8 together with vaccine set $V = \{1 = \{1,2\}, 2 = \{1,2,3\}\}$, disease set $D_{NE} = \{1,2\}$, and weights $\rho_d = \gamma_d = 1$ for diseases $d = 1,2$.

<table>
<tr><th rowspan="2">DISEASE</th><th colspan="9">TIME PERIOD</th></tr>
<tr><th>1</th><th>2</th><th>3</th><th>4</th><th>5</th><th>6</th><th>7</th><th>8</th><th>9</th></tr>
<tr><td>1</td><td colspan="3">Dose 1</td><td colspan="2">Dose 2</td><td></td><td></td><td colspan="2">Dose 3</td></tr>
<tr><td>2</td><td></td><td colspan="3">Dose 1</td><td>Dose 2</td><td></td><td colspan="2">Dose 3</td><td></td></tr>
<tr><td>3</td><td colspan="2">Dose 1</td><td></td><td></td><td colspan="2">Dose 2</td><td></td><td></td><td>Dose 3</td></tr>
</table>

Figure 8: Childhood Immunization Schedule for Example 12

Here,

$\mathbf{D}$ = {(1,1), (1,2), (1,3), (2,1), (2,2), (2,3), (3,1), (3,2), (3,3)}, and

C_{11} = {(1,1)}, C_{12} = {(1,1), (3,1)}, C_{21} = {(1,1), (2,1)}, C_{22} = {(1,1), (2,1), (3,1)},

C_{31} = {(1,1), (2,1)}, C_{32} = {(1,1), (2,1)}, C_{41} = {(1,2), (2,1)}, C_{42} = {(1,2), (2,1)},

C_{51} = {(1,2), (2,2)}, C_{52} = {(1,2), (2,2), (3,2)}, $C_{61} = \emptyset$, C_{62} = {(3,2)}, C_{71} = {(2,3)},

C_{72} = {(2,3)}, C_{81} = {(1,3),(2,3)}, C_{82} = {(1,3),(2,3)}, C_{91} = {(1,3)}, C_{92} = {(1,3), (3,3)}.

The *Greedy* heuristic proceeds as follows:

Step 1: Initialize:

a. $X_{tv} = 0$ for $t = 1,2,\ldots,9$ and $v = 1,2$

b. $\hat{C}_{tv} = C_{tv}$ for $t = 1,2,\ldots,9$ and $v = 1,2$

Step 2(1): $\mathbf{C} = \emptyset$ since $X_{tv} = 0$ for all $t \in T$, $v \in V$, and hence for non-empty sets $\hat{C}_{tv}$

a. $W_{11} = w_{11} + w_{21} = 0 + \gamma_2 = 1$, $W_{12} = w_{11} + w_{21} + w_{31} = 0 + \gamma_2 + 0 = 1$,

$W_{21} = w_{12} + w_{22} = 0 + 0 = 0$, $W_{22} = w_{12} + w_{22} + w_{32} = 0 + 0 + 0 = 0$,

$W_{31} = w_{13} + w_{23} = 0 + 0 = 0$, $W_{32} = w_{13} + w_{23} + w_{33} = 0 + 0 + 0 = 0$,

$W_{41} = w_{14} + w_{24} = 0 + 0 = 0$, $W_{42} = w_{14} + w_{24} + w_{34} = 0 + 0 + 0 = 0$,

$W_{51} = w_{15} + w_{25} = 0 + 0 = 0$, $W_{52} = w_{15} + w_{25} + w_{35} = 0 + 0 + 0 = 0$,

$W_{62} = w_{16} + w_{26} + w_{36} = \gamma_1 + \gamma_2 + 0 = 2$,

$W_{71} = w_{17} + w_{27} = \gamma_1 + 0 = 1$, $W_{72} = w_{17} + w_{27} + w_{37} = \gamma_1 + 0 + 0 = 1$,

$W_{81} = w_{18} + w_{28} = 0 + 0 = 0$, $W_{82} = w_{18} + w_{28} + w_{38} = 0 + 0 + 0 = 0$,

$W_{91} = w_{19} + w_{29} = 0 + \gamma_2 = 1$, $W_{92} = w_{19} + w_{29} + w_{39} = 0 + \gamma_2 + 0 = 1$,

b. $(t', v') = \underset{t \in T, v \in V}{\arg\min}\, W_{tv} / |\hat{C}_{tv}| = (2,2)$ or $(5,2)$, since $\hat{C}_{22}$ and $\hat{C}_{52}$ have the minimum weight and cover the most disease doses. Therefore, let

c. $X_{22} = 1$, which implies $\hat{C}_{tv} = \emptyset$ for $t = 1,2,3$ and $v = 1,2$ with remaining non-empty sets:

d. $\hat{C}_{41} = \{(1,2)\}$, $\hat{C}_{42} = \{(1,2)\}$, $\hat{C}_{51} = \{(1,2), (2,2)\}$, $\hat{C}_{52} = \{(1,2), (2,2), (3,2)\}$, $\hat{C}_{62} = \{(3,2)\}$, $\hat{C}_{71} = \{(2,3)\}$, $\hat{C}_{72} = \{(2,3)\}$, $\hat{C}_{81} = \{(1,3), (2,3)\}$, $\hat{C}_{82} = \{(1,3), (2,3)\}$, $\hat{C}_{91} = \{(1,3)\}$, $\hat{C}_{92} = \{(1,3), (3,3)\}$.

Step 2(2): $\mathbf{C} = \{(1,1), (2,1), (3,1)\} \neq \mathbf{D}$

a. $W_{41} = \rho_2 = 1$, $W_{42} = \rho_2 = 1$, $W_{51} = 0$, $W_{52} = 0$, $W_{62} = \gamma_1 + \gamma_2 = 2$,

$W_{71} = \gamma_1 = 1$, $W_{72} = \gamma_1 = 1$, $W_{81} = 0$, $W_{82} = 0$, $W_{91} = \gamma_2 = 1$, $W_{92} = \gamma_2 = 1$

b. $(t', v') = \underset{t \in T, v \in V}{\arg\min}\, W_{tv} / |\hat{C}_{tv}| = (5,2)$

c. $X_{52} = 1$

d. $\hat{C}_{71} = \{(2,3)\}$, $\hat{C}_{72} = \{(2,3)\}$, $\hat{C}_{81} = \{(1,3), (2,3)\}$, $\hat{C}_{82} = \{(1,3), (2,3)\}$, $\hat{C}_{91} = \{(1,3)\}$, $\hat{C}_{92} = \{(1,3), (3,3)\}$

Step 2(3): $\mathbf{C} = \{(1,1), (1,2), (2,1), (2,2), (3,1), (3,2)\} \neq \mathbf{D}$

a. $W_{71} = \gamma_1 = 1$, $W_{72} = \gamma_1 = 1$, $W_{81} = 0$, $W_{82} = 0$, $W_{91} = \gamma_2 = 1$, $W_{92} = \gamma_2 = 1$

b. $(t', v') = \underset{t \in T, v \in V}{\arg\min}\ W_{tv} / |\hat{C}_{tv}| = (8,1)$

c. $X_{81} = 1$

d. $\hat{C}_{92} = \{(3,3)\}$

Step 2(4): $\mathbf{C} = \{(1,1), (1,2), (1,3), (2,1), (2,2), (2,3), (3,1), (3,2)\} \neq \mathbf{D}$

a. $W_{92} = \rho_1 + \gamma_2 = 2$

b. $(t', v') = \underset{t \in T, v \in V}{\arg\min}\ W_{tv} / |\hat{C}_{tv}| = (9,2)$

c. $X_{92} = 1$

d. $\hat{C}_{tv} = \emptyset$ for $t = 1,2,\ldots,9$ and $v = 1,2$

STOP while loop, since $\mathbf{C} = \{(1,1), (1,2), (1,3), (2,1), (2,2), (2,3), (3,1), (3,2), (3,3)\} = \mathbf{D}$.

Step 3: $Z_{11}^{P} = 0,\ Z_{12}^{P} = 0,\ Z_{13}^{P} = 1,\ Z_{21}^{P} = 0,\ Z_{22}^{P} = 0,\ Z_{23}^{P} = 0,\ Z_{1}^{R} = 0$, and $Z_{2}^{R} = 1$

Step 4: Return $\sum_{d \in D_{NE}} \left(\rho_d \left(\sum_{j=1}^{n_d} Z_{dj}^{P}\right) + \gamma_d Z_d^{R}\right) = 2.$ □

The *Greedy* heuristic executes in $O(|\mathbf{D}|\tau\upsilon)$ time, and returns a feasible solution, since every iteration of the while loop (i.e., Step 2) administers a vaccine that satisfies at least one dose requirement for some disease $d \in D$ (i.e., every iteration covers at least one $(d, j) \in \mathbf{D}$). Therefore, the *Greedy* heuristic should (in practice) be more efficient than the *MAX Rounding* heuristic; however, the *Greedy* heuristic is not an approximation algorithm. To see this, consider the following instance of VFSREP(O)-MED: $T = \{1\}$, $D = D_{NE} = \{1,2,\ldots,2m\}$, and $V = \{1 = \{1,m+1\}, 2 = \{2,m+2\}, \ldots, m = \{m,2m\}, m+1 = \{1,2,\ldots,m\}\}$. Therefore, vaccines 1 through m are bivalent vaccines such that $I_{vd} = 1$ for vaccine $v \in V$ and diseases $d = v$ and $d = m + v$, 0 otherwise, and vaccine $v = m+1$ is a multivalent vaccine such that $I_{(m+1)d} = 1$ for diseases $d = 1,2,\ldots,m$, 0 otherwise. Since $\tau = 1$, then $n_d = 1$ for all diseases $d \in D$, and hence, $P_{djt} = 1$ for all diseases $d \in D$, $j = 1$, $t = 1$. Furthermore, suppose $\rho_d = \gamma_d = 1$ for all diseases $d \in D_{NE}$. Observe that the optimal solution is to administer the m bivalent vaccines, which would result in no extraimmunization. However, in Step 2.a of the *Greedy* heuristic, $W_{tv} = 0$ for time period $t = 1$ and for all vaccines $v \in V$, and in Step 2.b, $(t', v') = (1,m+1)$, since vaccine $v = m+1$ immunizes against the largest number of diseases in time period $t = 1$. Therefore, $X_{1(m+1)} = 1$ in Step 2.c of the *Greedy* heuristic. After this iteration of the while loop in Step 2, $\mathbf{C} = \{(1,1), (2,1), \ldots, (m,1)\}$, and hence, diseases $d = m+1, m+2, \ldots, 2m$ must still be satisfied by some vaccine. This implies that the m bivalent vaccines must also be selected by the *Greedy* heuristic in subsequent iterations. Therefore, the objective function value returned by the *Greedy* heuristic in Step 4 is $\sum_{d \in D_{NE}} \left(\rho_d \left(\sum_{j=1}^{n_d} Z_{dj}^P \right) + \gamma_d Z_d^R \right) =$ m, and hence, $\sum_{d \in D_{NE}} \left(\rho_d \left(\sum_{j=1}^{n_d} Z_{dj}^P \right) + \gamma_d Z_d^R \right) \to \infty$ as $m \to \infty$. Although the *Greedy* heuristic can be arbitrarily bad, the computational results in Section 5 show that the *Greedy* heuristic performed as well as the *MAX Rounding* heuristic for several randomly generated VFSREP(O)-MED instances.

4.3.4 Randomized Rounding Heuristic

This section presents the *Randomized Rounding* heuristics for VFSREP(O)-MED(A). The *Randomized Rounding* heuristic is shown to be a randomized approximation algorithm, which by definition, executes in polynomial time and provides an approximation bound on the expected value of the heuristic solution (Hochbaum 1997).

The *Randomized Rounding* heuristic uses the solution from an LP relaxation to construct a feasible binary solution for VFSREP(O)-MED(A). Relaxing the binary constraints for the decision variables in VFSREP(O)-MED(A) yields the LP relaxation

$$\text{Maximize} \quad \sum_{d \in D} \sum_{j=1}^{n_d} \pi_{dj}$$

Subject to

$$\sum_{t \in T} \sum_{v \in V} P_{djt} X_{tv} I_{vd} \geq \pi_{dj} \qquad \text{for all } d \in D_E,\ j = 1,2,\ldots,n_d,$$

$$\sum_{t \in T} \sum_{v \in V} P_{djt} X_{tv} I_{vd} = \pi_{dj} \qquad \text{for all } d \in D_{NE},\ j = 1,2,\ldots,n_d,$$

$$\sum_{t \in T} \sum_{v \in V} R_{dt} X_{tv} I_{vd} = 0 \qquad \text{for all } d \in D_{NE},$$

$$0 \leq X_{tv} \leq 1 \qquad \text{for all } t \in T,\ v \in V,$$

$$0 \leq \pi_{dj} \leq 1 \qquad \text{for all } d \in D,\ j = 1,2,\ldots,n_d.$$

Denote the optimal objective function values of VFSREP(O)-MED(A) and its LP relaxation as $z_{IP(A)}$ and $z_{LP(A)}$, respectively, where $z_{LP(A)} \geq z_{IP(A)}$ (since the feasible region of VFSREP(O)-MED(A) is contained in the feasible region of its LP relaxation). Let $X_{LP(A)}^*$ denote the optimal decision vector for the LP relaxation and $X_{LP(A)_{tv}}^*$, $t \in T$, $v \in V$, and π_{dj}^*, $d \in D$, j $=1,2,\ldots,n_d$, denote the optimal values for the decision variables in the LP relaxation. After solving the LP relaxation of VFSREP(O)-MED(A), the *Randomized Rounding* heuristic assigns binary decision variable X_{tv} = 1(0) with probability $X_{LP(A)_{tv}}^*$ (1 – $X_{LP(A)_{tv}}^*$) for each time period $t \in T$ and vaccine $v \in V$. This binary variable assignment is then used to determine the number of satisfied doses. The *Randomized Rounding* heuristic is now formally given.

Randomized Rounding Heuristic for VFSREP(O)-MED(A)

Step 1. Solve the LP relaxation of VFSREP(O)-MED(A)
Step 2. $X_{tv} \leftarrow 0$ for all $t \in T$ and $v \in V$
Step 3. $\pi_{dj} \leftarrow 0$ for all $d \in D$, $j = 1,2,\ldots,n_d$ and λ_d
Step 4. For $i = 1,2,\ldots,K$ ($K = 500$ for the computational results in Section 4.4)
- a. For all $t \in T$ and $v \in V$
  - i. Draw a random number *RAND*, where $RAND \sim U(0,1)$
  - ii. If $X^*_{LP(A)_{tv}} \geq RAND$, then $X_{tv} \leftarrow 1$
- b. For all $d \in D_{NE}$
  - i. For all $j = 1,2,\ldots,n_d$
    1. $\hat{\pi}_{dj} \leftarrow \Sigma_{t\in T}\Sigma_{v\in V} P_{djt} X_{tv} I_{vd}$
    2. If $d \in D_{NE}$ and $\hat{\pi}_{dj} > 1$, then set $X_{tv} \leftarrow 0$ such that $X_{tv} = 1$ and $P_{dt} = I_{vd} = 1$ for the $\hat{\pi}_{dj} - 1$ variables with the smallest fractional values $X^*_{LP(A)_{tv}}$
- c. For all $d \in D$
  - i. Compute $\hat{\pi}_{dj} \leftarrow \Sigma_{t\in T}\Sigma_{v\in V} P_{djt} X_{tv} I_{vd}$
    1. If $d \in D_{NE}$ and $\hat{\pi}_{dj} = 1$, then $\pi_{dj} \leftarrow 1$
    2. If $d \in D_E$ and $\hat{\pi}_{dj} \geq 1$, then $\pi_{dj} \leftarrow 1$
- d. Compute $SatisfiedDoses(i) = \sum_{d\in D}\sum_{j=1}^{n_d}\pi_{dj}$ for replication i

Step 5. Return $\max_i SatisfiedDoses(i)$

Example 13

Consider the childhood immunization schedule displayed in Figure 3 together with the vaccine set $V = \{1 = \{1\}, 2 = \{2\}, 3 = \{2,3\}\}$, disease set $D_{NE} = \{1,2\}$, weights $\rho_d = \gamma_d = 1$ for diseases $d = 1,2$, and suppose $K = 2$ (for Step 4 of the *Randomized Rounding* heuristic). The specific instance for VFSREP(O)-MED(A) for this example is:

Maximize $\pi_{11} + \pi_{12} + \pi_{21} + \pi_{22} + \pi_{31}$

Subject to

$X_{11} + X_{21} = \pi_{11}$ for $d = 1$, $j = 1$

$X_{31} + X_{41} = \pi_{12}$ for $d = 1$, $j = 2$

$X_{12} + X_{13} = \pi_{21}$ for $d = 2$, $j = 1$

$X_{42} + X_{43} = \pi_{22}$ for $d = 2, j = 2$

$X_{23} + X_{33} \geq \pi_{31}$ for $d = 3, j = 1$

$X_{22} + X_{23} + X_{32} + X_{33} = 0$ for $d = 2$

$0 \leq X_{tv} \leq 1$ for all $t \in T, v \in V$,

$0 \leq \pi_{dj} \leq 1$ for all $d \in D, j = 1,2,\ldots,n_d$.

In this example, $z_{LP(A)} = 4$, since a dose of vaccine $v = 3$ can not be administered in time period $t = 2$ or 3 to satisfy the requirement for disease $d = 3$ dose $j = 1$. Therefore, a feasible decision vector for $X^*_{LP} = (X^*_{LP_{11}}, X^*_{LP_{12}}, X^*_{LP_{13}}, X^*_{LP_{21}}, X^*_{LP_{22}}, X^*_{LP_{23}}, X^*_{LP_{31}}, X^*_{LP_{32}}, X^*_{LP_{33}}, X^*_{LP_{41}}, X^*_{LP_{42}}, X^*_{LP_{43}})$ that yields this optimal value is $X^*_{LP} = (1, \frac{5}{6}, \frac{1}{6}, 0,0,0, \frac{3}{4}, 0,0, \frac{1}{4}, \frac{2}{3}, \frac{1}{3})$. Let **R** represent the random numbers drawn in Step 4.a.i, where the k^{th} component of **R** is the random number that corresponds to the k^{th} component of the vector X^*_{LP}.

In Step 4, for $i = 1$, suppose **R** = (0.06, 0.35, 0.81, 0.01, 0.2, 0.79, 0.6, 0.27, 0.92, 0.44, 0.75, 0.89), then these random numbers yield the binary assignment (1,1,0,0,0,0,1,0,0,0,0,0) in Step 4.a.ii, which returns an objective function value of 3, since $\pi_{11} = \pi_{12} = \pi_{21} = 1$. Therefore, in Step 4.d, *SatisfiedDoses*(1) = 3.

In Step 4, for $i = 2$, suppose **R** = (0.82, 0.44, 0.11, 0.62, 0.92, 0.74, 0.41, 0.94, 0.39, 0.08, 0.57, 0.68), then these random numbers yield the binary assignment (1,1,1,0,0,0,1,0,0,1, 1,0) in Step 4.a.ii. In Step 4.b.i.1, $\hat{\pi}_{12} = \hat{\pi}_{21} = 2$, which is infeasible since disease $d = 1,2 \in D_{NE}$. Therefore, for disease $d = 1$, dose $j = 2$, either X_{31} or X_{41} must be set to zero, and for disease $d = 2$, dose $j = 1$, either X_{12} or X_{13} must be set to zero. In Step 4.b.i.2, $X_{41} = 0$ since $X^*_{LP_{31}} > X^*_{LP_{41}}$ and $X_{13} = 0$ since $X^*_{LP_{12}} > X^*_{LP_{13}}$, yielding the resulting binary assignment (1,1,1,0,0,0,1,0,0,0,1,0). This binary assignment implies $\pi_{11} = \pi_{12} = \pi_{21} = \pi_{22} = 1$, which returns the optimal objective function value of 4. Therefore, in Step 4.d, *SatisfiedDoses*(2) = 4, and hence, Step 5 returns $\max_i$ *SatisfiedDoses*(i) = max{3,4} = 4. □

The *Randomized Rounding* heuristic executes in $O(\mathbf{T}_{\text{LP}} + K(\tau^2 \upsilon\, \delta))$ time, where $\mathbf{T}_{\text{LP}}$ is the time required to solve the LP relaxation of VFSREP(O)-MED(A). Steps 4.b and 4.c ensure that the solution returned by the *Randomized Rounding* heuristic is feasible. Given that linear programming is solvable in polynomial time, it then follows that the *Randomized Rounding* heuristic executes in polynomial time. Theorem 11 shows that the expected value of the solution returned by the *Randomized Rounding* heuristic is guaranteed to be no worse than $\xi \cdot z_{IP(A)}$, where $\xi = (1 - p_{mzx})^{\alpha-1}$ for $p_{max} = \max_{t \in T, v \in V: 0 < X^*_{LP_{tv}} < 1} \{X^*_{LP_{tv}}\}$ and $\alpha = \max_{d \in D} (\Sigma_{v \in V} I_{vd})(\max_{j=1,2,\ldots,n_d} \Sigma_{t \in T} P_{djt})$ as defined above.

THEOREM 11: *The Randomized Rounding heuristic is a randomized ξ-approximation algorithm for VFSREP(O)–MED(A).*

PROOF: A randomized approximation algorithm is an algorithm that executes in polynomial time and guarantees the expected value of its returned solution is within some constant factor ξ (i.e., *Randomized Rounding* is a randomized ξ-approximation algorithm if $E[\mathbf{z}] \geq \xi \cdot z_{IP(A)}$, where **z** is the value of the objective function returned by the *Randomized Rounding* heuristic and $z_{IP(A)}$ is the optimal value for VFSREP(O)-MED(A)). Clearly, the *Randomized Rounding* heuristic executes in polynomial time (i.e., $O(\mathbf{T}_{\text{LP}} + K(\tau^2 \upsilon \ \delta))$) given that linear programming is solvable in polynomial time. For purposes of the approximation bound, assume $K = 1$ and define the following random variable

$$\hat{\pi}_{dj} = \begin{cases} 1 & \text{if disease } d \in D, \text{dose } j = 1,2,\ldots,n_d \text{ is satisfied} \\ 0 & \text{otherwise.} \end{cases}$$

Therefore, $\hat{\pi}_{dj} = 1$ if and only if $\Sigma_{t \in T} \Sigma_{v \in V} P_{djt} X_{tv} I_{vd} \geq 1$ for disease $d \in D_E$, dose $j = 1,2,\ldots, n_d$ and $\Sigma_{t \in T} \Sigma_{v \in V} P_{djt} X_{tv} I_{vd} = 1$ for disease $d \in D_{NE}$, dose $j = 1,2,\ldots,n_d$. Furthermore, the objective function value returned by the *Randomized Rounding* heuristic is $\mathbf{z} = \sum_{d \in D} \sum_{j=1}^{n_d} \hat{\pi}_{dj}$, and hence,

$$E[\mathbf{z}] = E[\sum_{d \in D} \sum_{j=1}^{n_d} \hat{\pi}_{dj}] = \sum_{d \in D} \sum_{j=1}^{n_d} E[\hat{\pi}_{dj}] = \sum_{d \in D} \sum_{j=1}^{n_d} (0 \cdot P[\hat{\pi}_{dj} = 0] + 1 \cdot P[\hat{\pi}_{dj} = 1]) = \sum_{d \in D} \sum_{j=1}^{n_d} P[\hat{\pi}_{dj} = 1].$$

Observe that $P[\hat{\pi}_{dj} = 1] = P[\Sigma_{t \in T} \Sigma_{v \in V} P_{djt} X_{tv} I_{vd} \geq 1]$ for disease $d \in D_E$, dose $j = 1,2,\ldots,n_d$ and $P[\hat{\pi}_{dj} = 1] = P[\Sigma_{t \in T} \Sigma_{v \in V} P_{djt} X_{tv} I_{vd} = 1]$ for disease $d \in D_{NE}$, dose $j = 1,2,\ldots,n_d$. Therefore, for disease $d \in D_E$, dose $j = 1,2,\ldots,n_d$,

$$P[\hat{\pi}_{dj} = 1] = P[\text{at least one } X_{tv} = 1] = 1 - P[\text{all } X_{tv} = 0],$$

and for disease $d \in D_{NE}$, dose $j = 1,2,\ldots,n_d$,

$$P[\hat{\pi}_{dj} = 1] = P[\text{exactly one } X_{tv} = 1] = 1 - (P[\text{more than one } X_{tv} = 1] + P[\text{all } X_{tv} = 0]).$$

Furthermore, P[at least one $X_{tv} = 1$] $\geq$ P[exactly one $X_{tv} = 1$] for all diseases $d \in D$ and doses $j = 1,2,\ldots,n_d$, and hence, for purposes of the approximation bound, assume that $D_E = \emptyset$. (It can be shown that the *Randomized Rounding* heuristic is a randomized (1-1/e)-approximation algorithm for VFSREP(O)-MED(A) when $D_{NE} = \emptyset$ by using similar arguments from MAX-SAT (Hochbaum 1997).)

Consider some disease $d \in D$ and dose $j = 1,2,\ldots,n_d$, then the LP relaxation of VFSREP(O)-MED(A) implies $\pi^*_{dj} = \Sigma_{t\in T}\Sigma_{v\in V}\ P_{djt}\, X^*_{LP_{tv}}\ I_{vd}$. Furthermore, for this constraint, suppose there are k fractional variables such that $0 < X^*_{LP_{tv}} < 1$. (If $X^*_{LP_{tv}} = 1$ for some $t \in T$ and $v \in V$, then $P[\hat{\pi}_{dj} = 1] = 1$, and if $X^*_{LP_{tv}} = 0$ for all $t \in T$ and $v \in V$, then $P[\hat{\pi}_{dj} = 1] = 0$). Then,

$P[\hat{\pi}_{dj} = 1] = P$[exactly one $X_{tv} = 1$] $= X^*_{LP_{tv1}}(1 - X^*_{LP_{tv2}})\,(1 - X^*_{LP_{tv3}}) \cdot\cdot\cdot (1 - X^*_{LP_{tvk}}) + (1 - X^*_{LP_{tv1}})\, X^*_{LP_{tv2}}(1 - X^*_{LP_{tv3}}) \cdot\cdot\cdot (1 - X^*_{LP_{tvk}}) + \cdot\cdot\cdot + (1 - X^*_{LP_{tv1}})(1 - X^*_{LP_{tv2}}) \cdot\cdot\cdot (1 - X^*_{LP_{tv(k-1)}})\, X^*_{LP_{tvk}}$,

which implies,

$$P[\hat{\pi}_{dj} = 1] \geq X^*_{LP_{tv1}}(1 - p)^{k-1} + X^*_{LP_{tv2}}(1 - p)^{k-1} + \cdot\cdot\cdot + X^*_{LP_{tvk}}(1 - p)^{k-1}, \text{ where } p = \max_{i=1,2,\ldots,k} X^*_{LP_{tvi}}$$

$$= (1 - p)^{k-1}\,(X^*_{LP_{tv1}} + X^*_{LP_{tv2}} + \cdot\cdot\cdot + X^*_{LP_{tvk}}) = (1 - p)^{k-1}\ \pi^*_{dj}.$$

Therefore,

$$E[\mathbf{z}] = \sum_{d\in D}\sum_{j=1}^{n_d} P[\hat{\pi}_{dj} = 1]$$

$$\geq \max_k\,(1 - p_{max})^{(k-1)} \sum_{d\in D}\sum_{j=1}^{n_d} \pi^*_{dj}\,, \text{ where } p_{max} = \max_{t\in T,\, v\in V: 0 < X^*_{LP_{tv}} < 1} \{X^*_{LP_{tv}}\}$$

$$\geq (1 - p_{max})^{(\alpha-1)} \cdot z_{LP(A)}, \text{ where } \alpha = \max_{d\in D} \alpha_d \text{ and } \alpha_d = (\Sigma_{v\in V}\, I_{vd})(\max_{j=1,2,\ldots,n_d} \Sigma_{t\in T}\, P_{djt})$$

$$\geq (1 - p_{max})^{(\alpha-1)} \cdot z_{IP(A)} \ \text{ (since } z_{LP(A)} \geq z_{IP(A)}\text{), and hence,}$$

$\xi = (1 - p_{max})^{(\alpha-1)}$ for the *Randomized Rounding* heuristic. ■

In the worst case, Theorem 11 shows that the *Randomized Rounding* heuristic could return a solution that is arbitrarily close to zero; however, the computational results in Section 4.4 show that the *Randomized Rounding* heuristic performed well in practice for several randomly generated VFSREP(O)-MED(A) instances.

4.3.5 Generalized Heuristic

This section presents the *MAX Rounding* and *Greedy* heuristics for VFSREP(O) by converting a VFSREP(O) instance into two distinct VFSREP(O)-MED instances, and then applying the *MAX Rounding* and *Greedy* heuristics for each VFSREP(O)-MED instance to find a feasible solution for the VFSREP(O) instance.

The *MAX Rounding* and *Greedy* heuristics for VFSREP(O)-MED do not ensure a feasible solution for an arbitrary VFSREP(O) instance, where some diseases $d \in D$ in the childhood immunization schedule do not have mutually exclusive doses. The reason these heuristics do not ensure feasibility for an arbitrary VFSREP(O) is because the sets C_{tv}, for all time periods $t \in T$ and vaccines $v \in V$, defined in the *MAX Rounding* and *Greedy* heuristics for VFSREP(O)-MED no longer satisfy unique dose requirements, since, for the diseases $d \in D$ that do not have mutually exclusive doses, there are time periods $t \in T$ when more than one required dose may be administered. For example, if vaccine $v \in V$ is a monovalent vaccine such that $I_{vd} = 1$ for disease $d \in D$, and in time period $t \in T$, $P_{djt} = 1(0)$ for $j = 1,2,(3,4,...,n_d)$ then $C_{tv} = \{(d,1),(d,2)\}$, and hence, administering vaccine $v \in V$ in time period $t \in T$ satisfies doses 1 and 2 for disease d. Therefore, to ensure the sets C_{tv} satisfy unique dose requirements, consider two variations of the set C_{tv} for all time periods $t \in T$ and vaccines $v \in V$

1) *Minimum Dose*: $C_{tv}^{MIN} = \{(d, k) \in \mathbf{D}: I_{vd} = 1 \text{ and } k = \min\{j: P_{djt} = 1\}\}$
2) *Maximum Dose*: $C_{tv}^{MAX} = \{(d, k) \in \mathbf{D}: I_{vd} = 1 \text{ and } k = \max\{j: P_{djt} = 1\}\}$.

Variations 1) or 2) ensure that set C_{tv} (i.e., $C_{tv} = C_{tv}^{MIN}$ for all time periods $t \in T$ and vaccines $v \in V$, or $C_{tv} = C_{tv}^{MAX}$ for all time periods $t \in T$ and vaccines $v \in V$) satisfies unique dose requirements for all diseases $d \in D$, and hence, each variation converts a VFSREP(O) instance into a distinct VFSREP(O)-MED instance.

Therefore, the *A* heuristic for VFSREP(O), where *A* is the *MAX Rounding* or *Greedy* heuristic, executes the *A* heuristic for VFSREP(O)-MED on each distinct VFSREP(O)-MED instance. The *A* heuristic is now formally given.

A Heuristic for VFSREP(O)

Step 1. Select $A \in \{$*MAX Rounding*, *Greedy*$\}$
Step 2. Initialize:
a. Let $\mathbf{D} = \{(d, j) : d \in D, j = 1,2,\ldots,n_d\}$ and $C_{tv} = C_{tv}^{MIN}$ for all $t \in T, v \in V$

Step 3. Execute *A heuristic for VFSREP(O)-MED* and return

$$Z_{MIN} = \sum_{d \in D_{NE}} \left(\rho_d \left(\sum_{j=1}^{n_d} Z_{dj}^P \right) + \gamma_d Z_d^R \right)$$

Step 4. Initialize:

a. Let $\mathbf{D} = \{(d, j) : d \in D, j = 1,2,\ldots,n_d\}$ and $C_{tv} = C_{tv}^{MAX}$ for all $t \in T, v \in V$

Step 5. Execute *A heuristic for VFSREP(O)-MED* and return

$$Z_{MAX} = \sum_{d \in D_{NE}} \left(\rho_d \left(\sum_{j=1}^{n_d} Z_{dj}^P \right) + \gamma_d Z_d^R \right)$$

Step 6. Return $\min\{Z_{MIN}, Z_{MAX}\}$

The *A* heuristic executes in $O(|\mathbf{D}|\tau\upsilon)$ time for A = *Greedy* and $O(\mathbf{T}_{\text{LP}} + |\mathbf{D}|\tau\upsilon)$ time for A = *MAX Rounding*, where $\mathbf{T}_{\text{LP}}$ is the time required to solve the LP relaxations of both distinct VFSREP(O)-MED instances.

Furthermore, the *A* heuristic returns a feasible solution for VFSREP(O) provided that, for all diseases $d \in D$, dose $j = 1,2,\ldots,n_d$ does not dominate dose $k = 1,2,\ldots,n_d$, $j \neq k$. Recall that dose $j = 1,2,\ldots,n_d$ is said to *dominate* dose $k = 1,2,\ldots,n_d$, $j \neq k$, for disease $d \in D$ if $P_{djt} \geq P_{dkt}$ for all time periods $t \in T$. If disease $d \in D$ has no dominant doses, then the time periods when dose $j = 1,2,\ldots,n_d$ may be administered do not completely overlap with the time periods when dose $k = 1,2,\ldots,n_d$, $j \neq k$, may be administered, and hence, for all $j = 1,2,\ldots,(n_d - 1)$, there exists time periods $t, t' \in T$ such that $P_{djt} = 1$ and $P_{d(j+1)t} = 0$ and $P_{djt'} = 0$ and $P_{d(j+1)t'} = 1$.

All of the diseases in the 2006 Recommended Childhood Immunization Schedule do not have a dose that dominates any other dose, and future schedules should also meet this restriction, since there is a biological spacing requirement between each dose of vaccine for every disease $d \in D$, and hence, requiring that the childhood immunization schedule has no dominant doses ensures that every $(d, j) \in \mathbf{D}$ (in Steps 2.a and 4.a) is contained in some set C_{tv} for at least one time period $t \in T$ and vaccine $v \in V$. Therefore, the *A* heuristic returns a feasible solution for VFSREP(O) (assuming VFSREP(O) has a feasible solution), since every iteration of the *A* heuristic for VFSREP(O)-MED (in Step 3 and Step 5) administers a vaccine that satisfies at least one dose requirement for some disease $d \in D$ (i.e., every iteration covers at least one $(d, j) \in \mathbf{D}$).

4.4 Computational Results

This section reports computational results comparing the *MAX Rounding* and *Greedy* heuristics, and the DP algorithm for VFSREP(O)-MED, and the *Randomized Rounding* heuristic for VFSREP(O)-MED(A) . Computational results are also reported for an IP branch and bound (IP B&B) algorithm for VFSREP(O)-MED (denoted IP-MIN) and VFSREP(O)-MED(A) (denoted IP-MAX). Computational results with the *Rounding* heuristic are not reported, since this heuristic (empirically) yields significant amounts of extraimmunization. The *MAX Rounding*, *Greedy*, and *Randomized Rounding* heuristics and the DP and IP B&B algorithms were executed on three sets of test problems to demonstrate their computational effectiveness and limitations. The first test problem is the 2006 Recommended Childhood Immunization Schedule. The second set of test problems are randomly generated based on hypothetical near-term future childhood immunization schedules, while the third set of test problems are larger, randomly generated childhood immunization schedules executed with several different vaccine sets. The size of these randomly generated childhood immunization schedules assume that future Recommended Childhood Immunization Schedules will expand to include more diseases and time periods, and hence, will require a larger number of both monovalent and combination vaccines. These assumptions are reasonable, given recent trends in expanding the schedule. For example, four time periods and three diseases have been added to the Recommended Childhood Immunization Schedule since 1995, and there are currently several pediatric vaccine products under development (CDC 1995, Cochi 2005, Infectious Diseases in Children 2002).

For the *MAX Rounding* and *Greedy* heuristics and the DP and IP-MIN exact algorithms for VFSREP(O)-MED, the solution quality effectiveness measure **z** is the value of the objective function. Furthermore, for all childhood immunization schedules, $\rho_d = \gamma_d = 1$ for all diseases $d \in D_{NE}$, and hence, **z** specifies the number of extra vaccine doses administered. Moreover, for the *Randomized Rounding* heuristic and IP-MAX algorithm for VFSREP(O)-MED(A), the solution quality effectiveness measure $\boldsymbol{\zeta}$ is the value of the objective function divided by the total number of required doses, and hence, $\boldsymbol{\zeta} = \mathbf{z} / \Sigma_{d=1}^{\delta} n_d$. Therefore, $\boldsymbol{\zeta}$ represents the percentage of dose requirements satisfied by the *Randomized Rounding* heuristic or IP-MAX algorithm for VFSREP(O)-MED(A). The execution time (in CPU seconds) is also reported for each heuristic and exact algorithm, which is the efficiency effectiveness measure. All heuristics and exact algorithms were coded and executed in *MATLAB*v7.0 on a 2.4 MHz Pentium IV with 1GB of RAM including both IP B&B algorithms. The IP-MIN algorithm used an open source mixed integer optimization routine (see Tawfik 2005) and the IP-MAX algorithm used the existing binary solver in the *MATLAB* optimization toolbox.

The first test problem is the 2006 Recommended Childhood Immunization Schedule displayed in Figure 1. Therefore, D = {1 = Hepatitis B, 2 = Diphtheria-Tetanus-Pertussis, 3 = *Haemophilus influenzae* type b, 4 = Polio, 5 = Measles-Mumps-Rubella, 6 = Varicella, 7 = Pneumococcus, 8 = Influenza, 9 = Hepatitis A} with dose vector $n = (3, 5, 4, 4, 2, 1, 4, 1, 2)$, since diphtheria, tetanus, and pertussis are considered one disease and measles, mumps, and rubella are also considered one disease, and $T = \{1,2,\ldots,10\}$. The schedule parameters P_{djt}, Q_{dt}, and R_{dt}, for diseases $d \in D$, dose $j = 1,2,\ldots,n_d$, and time periods $t \in T$ are all obtained from Figure 1. For example, for disease d = 1 = Hepatitis B and dose j = 2, P_{djt} = 1(0) for time periods t = 2,3(1,4,5,6,7,8,9,10). Two different sets of vaccines, V_1 and V_2, are evaluated on two different sets of diseases that restrict extraimmunization, D_{NE1} and D_{NE2}. The vaccine sets are V_1 = {1 = {1}, 2 = {2}, 3 = {3}, 4 = {4}, 5 = {5}, 6 = {6}, 7 = {7}, 8 = {8}, 9 = {9}, 10 = {2,3}, 11 = {1,3}, 12 = {1,2,4}, 13 = {5,6}} and V_2 = {1 = {1}, 2 = {2,3,4}, 3 = {1,9}, 4 = {4}, 5 = {5}, 6 = {6}, 7 = {7}, 8 = {8}, 9 = {9} 11 = {1,3}, 12 = {1,2,4}, 13 = {5,6}, 14 = {1,2,3,4}}, where V_1 represents a set of pediatric vaccines currently licensed for use in the United States and V_2 represents a set of pediatric vaccine with fewer monovalent vaccines and more combination vaccines, some of which are not yet licensed for use in the United States, but are projected to be in the future. The parameters I_{vd} are indicated within the sets V_1 and V_2, respectively. For example, vaccine 1 $\in V_1$ is the monovalent vaccine for disease 1 (Hepatits B), vaccine 12 $\in V_1$ is the combination vaccine *Pediarix*® that immunizes against diseases 1 (Hepatitis B), 2 (Diphtheria-Tetanus-Pertussis), and 4 (Polio), and vaccine 2 $\in V_2$ is the combination vaccine *Pentacel*® that immunizes against diseases 2 (Diphtheria-Tetanus-Pertussis), 3 (*Haemophilus influenzae* type b), and 4 (Polio), which was recently submitted for licensing in the United States. The disease sets are D_{NE1} = {1,2,3,4} and $D_{NE2} = D$ = {1,2,3,4,5,6,7,8,9}. Table 6 reports the solution quality and execution time (in CPU seconds) for each heuristic and exact algorithm and for each vaccine set and disease set combination.

Table 6: Computational Results for 2006 Recommend Childhood Immunization Schedule

VFSREP(O)-MED	V_1 and D_{NE1}		V_1 and D_{NE2}		V_2 and D_{NE1}		V_2 and D_{NE2}	
	z	**Time**	**z**	**Time**	**z**	**Time**	**z**	**Time**
MAX Rounding	0	0.14	1	0.27	3	0.17	4	0.25
Greedy	0	0.09	0	0.11	1	0.09	1	0.11
DP	0	0.36	0	0.36	1	0.47	1	0.45
IP-MIN		9.66		5.02		16.94		12.89
VFSREP(O)-MED(A)	ζ	**Time**	ζ	**Time**	ζ	**Time**	ζ	**Time**
Randomized Rounding	1.00	0.56	1.00	0.75	0.96	0.56	0.96	0.81
IP-MAX	1.00	0.20	1.00	0.27	0.96	0.22	0.96	0.22

Lemma 6 and Theorem 9 in Section 4.2 showed that VFSREP(O) is solvable in polynomial time when all vaccines $v \in V$ are monovalent vaccines or when there exists a corresponding monovalent vaccine for every disease $d \in D$, and hence, the results for the solution quality and execution time reported in Table 6 are not surprising, given that most diseases have a corresponding monovalent vaccine (particularly in vaccine set V_1). Furthermore, the *MAX Rounding* and *Greedy* heuristics for VFSREP(O)-MED were both more efficient than the exact algorithms, and in all cases, the *Greedy* heuristic returned the optimal solution. Moreover, the *Randomized Rounding* heuristic for VFSREP(O)-MED(A) also returned the optimal solution for each case, but took more time to execute than the IP-MAX algorithm since the random rounding was replicated $K = 500$ times. Excluding the IP-MIN algorithm, all heuristics and exact algorithms executed in less than a second. However, as the next set of test problems will illustrate, this is unlikely to occur for future Recommended Childhood Immunization Schedules, as the schedule expands and more combination vaccines are licensed for use and enter the market.

The second set of test problems considers hypothetical near-term future childhood immunization schedules. Each heuristic and exact algorithm were executed on 30 randomly generated childhood immunization schedules with 15 time periods, 30 vaccines, and 11 diseases. Therefore, each random childhood immunization schedule reflects a gradual expansion in the sets D (from 9 to 11 diseases) and T (from 10 to 15 time periods) and a significant increase in the number of available vaccines, particularly, combination vaccines. In each random childhood immunization schedule, $1 \leq n_d \leq 5$ for all diseases $d \in D$, $1 \leq$

$Val(v) \leq 6$ for all vaccines $v \in V$, and $P_{djt} = 1$ for at most three time periods $t \in T$ for every disease $d \in D$ and dose $j = 1,2,\ldots,n_d$. For each randomly generated childhood immunization schedule, each heuristic and exact algorithm was executed three times, where in execution 1, 2, and 3, δ_{NE} = 4, 8, and 11, respectively. Table 7 reports the solution quality and execution time (in CPU seconds) averaged over the 30 random childhood immunization schedules for each value of δ_{NE}. An additional measure $\boldsymbol{\lambda}$ that indicates the number of childhood immunization schedules that the respective heuristic or exact algorithm found the optimal solution is also reported. The IP B&B algorithms found the optimal solution for $\boldsymbol{\lambda}$ of the 30 random childhood immunization schedules, but exceeded the default execution time limit (two hours) or default iteration limit (10^7) for the remaining (30 – $\boldsymbol{\lambda}$) random childhood immunization schedules. The statistics reported in Table 7 are averaged over the $\boldsymbol{\lambda}$ random childhood immunization schedules for which the IP B&B algorithms found the optimal solution, which is why the average $\mathbf{z}$ values for the IP-MIN and DP algorithms differ when δ_{NE} = 8 and 11. Furthermore, the *Randomized Rounding* heuristic found a feasible solution for all VFSREP(O)-MED(A) instances, however, for comparative purposes, the statistics reported in Table 7 are for the instances for which the IP-MAX algorithm also found the optimal solution.

Table 7: Computational Results for Future Childhood Immunization Schedule

VFSREP(O)-MED	$\delta_{NE} = 4$			$\delta_{NE} = 8$			$\delta_{NE} = 11$		
	z	**Time**	λ	**z**	**Time**	λ	**z**	**Time**	λ
MAX Rounding	0.70	0.80	25	7.53	0.83	4	16.77	0.85	0
Greedy	1.03	0.23	18	7.87	0.29	2	13.97	0.35	0
DP	0.43	1.74	30	4.77	1.80	30	10.33	1.81	30
IP-MIN	0.43	518	30	4.48	1095	27	10.07	1767	27
VFSREP(O)-MED(A)	ζ	**Time**	λ	ζ	**Time**	λ	ζ	**Time**	λ
Randomized Rounding	0.98	0.67	25	0.72	0.85	8	0.57	0.92	10
IP-MAX	0.98	1.06	30	0.77	3.91	29	0.59	7.48	25

The data reported in Table 7 show that for VFSREP(O)-MED the *MAX Rounding* and *Greedy* heuristics found better solutions for $\delta_{NE} << \delta$. For the smaller values of δ_{NE}, the *MAX Rounding* heuristic slightly outperformed the *Greedy* heuristic, while the *Greedy* heuristic outperformed the *MAX Rounding* heuristic for $\delta_{NE} = \delta$. Across all values of δ_{NE}, the *Greedy* heuristic was the most efficient compared with *MAX Rounding* and the DP and IP-MIN exact algorithms. The DP algorithm executed 2 to 7 times slower than the heuristics; however, the IP-MIN algorithm, on average, executed approximately 300 to 1000 times slower than the DP algorithm. Furthermore, the DP algorithm found the optimal solution for all 90 instances of VFSREP(O)-MED reported in Table 7, with little sensitivity to the value of δ_{NE}. Conversely, the IP-MIN algorithm only found the optimal solution for 84 of the 90 instances of VFSREP(O)-MED reported in Table 7, and the average execution time for IP-MIN algorithm more than tripled when δ_{NE} went from 4 diseases to 11 diseases. The data reported in Table 7 show that for VFSREP(O)-MED(A) the *Randomized Rounding* heuristic also found better solutions when $\delta_{NE} << \delta$, and, on average, was always more efficient than the exact IP-MAX algorithm. Furthermore, the IP-MAX algorithm shared similar trends with the IP-MIN algorithm in that its execution time and ability to find an optimal solution in a 'reasonable' amount of time and memory was sensitive to the value of δ_{NE}.

The observed difference in execution time between the heuristics and exact algorithms reported in Table 7 could become problematic for larger childhood immunization schedules and/or for practical uses, since the execution time of the DP algorithm rapidly grows as the size of the disease set D grows. For example, a webpage used to find a vaccine formulary for a given childhood immunization schedule that limits the amount of extraimmunization would require an algorithm to execute in real-time, since most web users would terminate a web application that required several seconds or minutes to execute. Moreover, the difference in execution time between the heuristics and exact algorithms will provide a more efficient analysis of larger childhood immunization schedules that may involve Monte Carlo simulation (see Jacobson and Sewell 2002) or the balking problem (described in Section 3.3.1), where either of these may require the solution of hundreds of thousands of VFSREP(O) instances. Furthermore, the childhood immunization schedule may need to be solved for each child, on a case-by-case basis, depending on the desired diseases in the set D_{NE}, and hence, efficient algorithms are needed so as to provide, in real-time, practical value for the public health community.

The third set of test problems considers larger randomly generated childhood immunization schedules that demonstrate the effect of combination vaccines and further demonstrate how a childhood immunization schedule's size affects the efficiency and solution quality of each heuristic and exact algorithm. Each heuristic and exact algorithm were executed on 30 randomly generated childhood immunization schedules with 20 time periods, 50 vaccines, and 13 diseases, where $\delta_{NE} = 10$, $1 \leq n_d \leq 5$ for all diseases $d \in D$, and $P_{djt} = 1$ for at most four time periods $t \in T$ for every disease $d \in D$ and dose $j = 1,2,\ldots,n_d$. For each randomly generated childhood immunization schedule, each heuristic and exact algorithm was executed six times, where for execution $i = 1,2,\ldots,6$, $Val(v) \leq i$ for all vaccines $v \in V$. Table 8 reports the solution quality and execution time (in CPU seconds) averaged across all 30 randomly generated childhood immunization schedules as well as the measure $\boldsymbol{\lambda}$ for each heuristic, algorithm, and valency. The overall average and standard deviation across all vaccine sets is also reported for each heuristic and exact algorithm. Note that computational results for the IP-MIN algorithm are not reported since the algorithm consistently exceeded virtual memory limits of 3.7 GB. Furthermore, the IP-MAX algorithm found the optimal solution for 151 of the 180 VFSREP(O)-MED(A) instances (each random childhood immunization schedule was executed six times), but exceeded the maximum number of iterations allowed for the remaining 29 instances (3 instances for $Val(v) \leq 4$, 11 instances for $Val(v) \leq 5$, and 15 instances for $Val(v) \leq 6$). The statistics reported in Table 8 are for the 151 instances for which the IP-MAX algorithm found the optimal solution while the execution time in parenthesis is the average over all instances. Moreover, the *Randomized Rounding* heuristic found a feasible solution for all 180 VFSREP(O)-MED(A) instances, however, for comparative purposes, the statistics reported in Table 8 are for the 151 instances for which the IP-MAX algorithm also found the optimal solution.

Table 8: Computational Results for the Effect of Combination Vaccines

	VFSREP(O)-MED								
	MAX Round			***Greedy***			**DP**		
***Val*(*v*) ≤**	**z**	**Time**	**λ**	**z**	**Time**	**λ**	**z**	**Time**	**λ**
1	0	1.1	30	0	1.3	30	0	6.4	30
2	0.8	0.9	14	1.1	0.9	18	0.4	9.1	30
3	2.3	1.0	8	3.4	0.8	6	1.6	11.2	30
4	5.0	1.2	3	4.6	0.8	1	2.1	12.8	30
5	6.7	1.3	1	6.2	0.7	2	3.1	14.4	30
6	8.5	1.5	0	7.9	0.7	1	4.6	15.7	30
Average	3.9	1.2	9.3	3.9	0.9	9.7	2.0	11.6	30
St Dev	3.4	0.2	11.4	3.0	0.2	11.9	1.7	3.5	0
	VFSREP(O)-MED(A)								
	Randomized Rounding				**IP-MAX**				

$Val(v) \leq$	ζ	Time	λ	ζ	Time	λ
1	1.00	1.1	30	1.00	1.1	30
2	0.99	1.2	28	0.99	1.3	30
3	0.92	1.2	16	0.94	2.7	30
4	0.88	1.3	7	0.92	8.6 (1184)	27
5	0.80	1.3	2	0.87	27.5 (4210)	19
6	0.68	1.3	1	0.77	8.9 (5957)	15
Average	0.88	1.2	14	0.91	8.4 (1893)	25
St Dev	0.11	0.1	12.8	0.09	10.0 (2574)	6.6

The data reported in Table 8 demonstrate how the size of the childhood immuninzation schedule and the complexity of the vaccine set impact the execution time and solution quality of the heuristics and exact algorithms. For example, in most cases, the execution time required to execute the heuristics and exact algorithms steadily increased as the valency (i.e., complexity) of the vaccine set increased. Furthermore, the IP-MIN algorithm only found the optimal solution for half of the random childhood immunization schedules when hexavalent vaccines were present. Moreover, the difference in execution times between the heuristics and exact algorithms significantly widened when compared to the results in Table 7. For example, the average execution time for the DP algorithm reported in Table 8 is more than 6 times the average execution time reported in Table 7, whereas the average execution times for the heuristics reported in Table 8 are less than 3 times the average execution times reported in Table 7. Observe also that the average solution quality for the *MAX Round* and *Greedy* heuristics were comparable across all vaccine sets, where the *MAX Round* heuristic performed better for lower valency vaccines and the *Greedy* heuristic performed better for higher valency vaccines. Furthermore, the execution time of the *Greedy* heuristic decreased as the valency of the vaccine set increased (almost requiring half the time when $Val(v) \leq 6$ compared when $Val(v) = 1$), which is intuitive, since this heuristic should require less iterations when more vaccines are able to protect against multiple diseases. As expected, the solution quality of all the heuristics deteriorated as the valency of the vaccine set increased. However, in the case of monovalent vaccines, each heuristic found the optimal solution on every random childhood immunization schedule, which is consistent in light of Theorem 9.

The computational results suggest that on average, the DP algorithm requires significantly less computational effort to find the optimal solution than is required by an IP B&B algorithm, such as IP-MIN, and with less variability. As shown in Section 4.3.1, the computational complexity of the DP algorithm is highly sensitive to the number of diseases, since the cardinality of the decision space is bounded above by 2^{δ} and $\mathbf{S}_{Max}$ also depends on the number of diseases. On the other hand, the computational complexity of the IP-MIN algorithm is highly sensitive to the number of decision variables, since the number of decision variables bound the number of possible branches. Furthermore, the computational effort of any IP B&B algorithm is sensitive to the gap between the value of the optimal integer solution and the corresponding value of the optimal LP relaxation solution, since a large gap would tend to require more branching to find the optimal integer solution (Nemhauser and Wolsey 1999). The data reported in Tables 6-8 provide empirical support for these remarks. For example, the DP algorithm was significantly more efficient than the IP-MIN algorithm reported in Table 7, and, on average, found the optimal solution for the larger childhood immunization schedules in less than 12 seconds versus the IP-MIN algorithm, which was unable to solve the larger childhood immunization schedules reported in Table 8 due to the amount of memory required to store the branch tree.

Chapter 5: The Complete Problem: Combining Budget and Extraimmunization

This chapter extends the models presented in Chapters 3 and 4 by presenting a general model that determines the set of vaccines that should be used in a clinical environment to satisfy any given childhood immunization schedule at minimum cost while also restricting extraimmunization. The chapter is organized as follows. Section 5.1 presents the general model (formulated as a decision problem and as a discrete optimization problem). Section 5.2 presents the computational complexity of the decision/discrete optimization problems and presents several formulation extensions. Lastly, Section 5.3 discusses how the DP algorithm and the *Rounding*, *MAX Rounding*, and *Greedy* heuristics for VFSREP(O) presented in Chapter 4 may be extended for the discrete optimization problem presented here.

5.1 Model Formulation

This section presents a model formulation for a decision problem and a discrete optimization problem used to design a vaccine formulary that addresses the cost of satisfying a given childhood immunization schedule while also restricting extraimmunization. Given a childhood immunization schedule, the decision problem, termed the Vaccine Formulary Selection with Limited Budget and Restricted Extraimmunization Problem (VFSLBREP), asks whether it is possible to design a vaccine formulary within a specified budget that also restricts extraimmunization for a specified set of diseases. This problem is now formally stated.

Vaccine Formulary Selection with Limited Budget and Restricted Extraimmunization Problem (VFSLBREP)

Given:

- A set of time periods, $T = \{1,2,\ldots,\tau\}$,
- a set of diseases, $D = \{1,2,\ldots,\delta\}$,
- a set of diseases where extraimmunization is permitted, $D_E \subseteq D$, with $|D_E| = \delta_E$,
- a set of diseases where extraimmunization is not permitted, $D_{NE} = D \backslash D_E$, with $|D_{NE}| = \delta_{NE}$,

- a set of vaccines $V = \{1,2,\ldots,\upsilon\}$, available to be administered to immunize against the δ diseases,
- the number of doses of a vaccine that must be administered for immunization against the δ diseases, $n_1,n_2,\ldots, n_\delta$,
- the cost of each vaccine, $c_1,c_2,\ldots,c_\upsilon$,
- a budget B,
- a set of binary parameters that indicate which vaccines immunize against which diseases; therefore, $I_{vd} = 1$ if vaccine $v \in V$ immunizes against disease $d \in D$, 0 otherwise,
- a set of binary parameters that indicate the set of time periods in which a particular dose of a vaccine may be administered to immunize against a disease; therefore, $P_{djt} = 1$ if in time period $t \in T$, a vaccine may be administered to satisfy the j^{th} dose, $j = 1,2,\ldots,n_d$, requirement for disease $d \in D$, 0 otherwise,
- a set of binary parameters that indicate the set of time periods in which a vaccine may be administered to satisfy any dose requirement against a disease; therefore, $Q_{dt} = 1$ if in time period $t \in T$, a vaccine may be administered to satisfy any dose requirement against disease $d \in D$, 0 otherwise, (i.e., for any disease $d \in D$ and time period $t \in T$, $Q_{dt} = 1$ if and only if $P_{djt} = 1$ for some dose $j = 1,2,\ldots,n_d$.),
- a set of binary parameters that indicate the set of time periods in which no dose of a vaccine may be administered to immunize against a disease where extraimmunization is not permitted; therefore, $R_{dt} = 1$ if in time period $t \in T$, no dose of a vaccine may be administered to immunize against disease $d \in D_{NE}$, 0 otherwise, (i.e., for any disease $d \in D_{NE}$ and time period $t \in T$, $R_{dt} = 1$ if and only if $Q_{dt} = 0$),
- a set of integer parameters that indicate the minimum number of doses of a vaccine required for disease $d \in D$ through time period $t \in T$; denoted m_{dt}.

Question: Does there exist a set of vaccines from V that can be administered over the time periods in T such that these vaccines immunize against all the diseases in D, at a total cost no greater than B while also restricting extraimmunization (i.e., do there exist values for the binary decision variables X_{tv}, $t \in T$, $v \in V$, where $X_{tv} = 1$ if vaccine $v \in V$ is administered in time period $t \in T$, 0 otherwise, and for the binary variables U_{dt}, $d \in D$, $t \in T$, where $U_{dt} = 1$ if any vaccine $v \in V$ that immunizes against disease $d \in D$ is administered in time period $t \in T$, 0 otherwise, such that for all diseases $d \in D$, $\Sigma_{t\in T}\Sigma_{v\in V}\, P_{djt}\, X_{tv}\, I_{vd} \geq 1$ for dose $j = 1,2,\ldots,n_d$, $\Sigma_{t=1,2,\ldots,t'}\, Q_{dt}\, U_{dt} \geq m_{dt'}$ for all time periods $t' \in T$, and $\Sigma_{v\in V}\, Q_{dt}\, X_{tv}\, I_{vd} \geq U_{dt}$ for all time periods $t \in T$, and for all diseases $d \in D_{NE}$, $\Sigma_{t\in T}\Sigma_{v\in V}\, R_{dt}\, X_{tv}\, I_{vd} = 0$ and $\Sigma_{t\in T}\, \Sigma_{v\in V}\, Q_{dt}\, X_{tv}\, I_{vd} = n_d$, and, finally, $\Sigma_{t\in T}\Sigma_{v\in V}\ c_v\, X_{tv} \leq B$)?

In the formulation of VFSLBREP, as was the case for VFSLBP and VFSREP, the given sets and parameters equate to a childhood immunization schedule together with budget and vaccine cost information. Likewise, all of the assumptions made for VFSLBP and VFSREP also apply to VFSLBREP. For example, the doses for all diseases $d \in D$ are assumed to be sequentially ordered. The question in VFSLBREP asks if there exists a vaccine formulary administered over the time periods in T that *satisfies* a given childhood immunization schedule and is within the given budget B and restricts extraimmunization for the diseases in the set D_{NE} (i.e., a variable assignment for the binary decision variables X_{tv}, for all time periods $t \in T$ and vaccines $v \in V$, and for the binary decision variables U_{dt}, for all diseases $d \in D$ and time periods $t \in T$, that satisfies the per dose requirements ($\Sigma_{t\in T}\Sigma_{v\in V}\, P_{djt}\, X_{tv}\, I_{vd} \geq 1$ for dose $j = 1,2,\ldots,n_d$) and total dosage requirements ($\Sigma_{t=1,2,\ldots,t'}\ Q_{dt}\ U_{dt} \geq m_{dt'}$ for all time period $t' \in T$ and $\Sigma_{v\in V}\, Q_{dt}\, X_{tv}\, I_{vd} \geq U_{dt}$ for all time periods $t \in T$) for each disease $d \in D$, and does not exceed the total dosage requirements ($\Sigma_{t\in T}\, \Sigma_{v\in V}\, Q_{dt}\, X_{tv}\, I_{vd} = n_d$) or provide a dose in a time period when no dose of a vaccine may be administered ($\Sigma_{t\in T}\Sigma_{v\in V}\, R_{dt}\, X_{tv}\, I_{vd} = 0$) for each disease $d \in D_{NE}$, and, finally, that satisfies the budget constraint ($\Sigma_{t\in T}\Sigma_{v\in V}\ c_v\, X_{tv} \leq B$)).

This decision problem can be addressed by solving a discrete optimization problem. More specifically, the following integer program can be used to answer VFSLBREP.

Integer Programming Model for Vaccine Formulary Selection with Limited Budget Problem (VFSLBREP(O))

$$\text{Minimize} \quad \sum_{t\in T}\sum_{v\in V} c_v X_{tv} + \sum_{d\in D_{NE}} \rho_d Z_d^P + \gamma_d Z_d^R \qquad \text{(O)}$$

Subject to

$$\sum_{t\in T}\sum_{v\in V} P_{djt} X_{tv} I_{vd} \geq 1 \qquad \text{for all } d \in D,\ j = 1,2,\ldots,n_d, \qquad (1)$$

$$\sum_{t=1,2,\ldots,t'} Q_{dt} U_{dt} \geq m_{dt'} \qquad \text{for all } d \in D,\ t' \in T, \qquad (2)$$

$$\sum_{v \in V} Q_{dt} X_{tv} I_{vd} \geq U_{dt} \qquad \text{for all } d \in D,\ t \in T, \qquad (3)$$

$$\sum_{t \in T} \sum_{v \in V} Q_{dt} X_{tv} I_{vd} - Z_d^P = n_d \qquad \text{for all } d \in D_{NE}, \qquad (4)$$

$$\sum_{t \in T} \sum_{v \in V} R_{dt} X_{tv} I_{vd} - Z_d^R = 0 \qquad \text{for all } d \in D_{NE}, \qquad (5)$$

$$X_{tv} \in \{0,1\} \qquad \text{for all } t \in T,\ v \in V, \qquad (6)$$

$$U_{dt} \in \{0,1\} \qquad \text{for all } d \in D,\ t \in T, \qquad (7)$$

$$Z_d^P,\ Z_d^R \in \mathbf{Z}^+ \cup \{0\} \qquad \text{for all } d \in D_{NE}, \qquad (8)$$

where sets T, D, D_{NE} and V, parameters $\{n_d\}$, $\{I_{vd}\}$, $\{P_{djt}\}$, $\{Q_{dt}\}$, $\{R_{dt}\}$, and $\{m_{dt}\}$, and variables $\{X_{tv}\}$ and $\{U_{dt}\}$ are defined in VFSLBREP, and

- $\rho_d \in \mathbf{Q}^+$ is the weight of extraimmunization for disease $d \in D_{NE}$ for all time periods $t \in T$ such that $Q_{dt} = 1$,
- $\gamma_d \in \mathbf{Q}^+$ is the weight of extraimmunization for disease $d \in D_{NE}$ for all time periods $t \in T$ such that $R_{dt} = 1$,
- $Z_d^P \in \mathbf{Z}^+ \cup \{0\}$ is the number of extra doses of vaccine administered for disease $d \in D_{NE}$ in all time periods $t \in T$ such that $Q_{dt} = 1$, and
- $Z_d^R \in \mathbf{Z}^+ \cup \{0\}$ is the number of extra doses of vaccine administered for disease $d \in D_{NE}$ in all time periods $t \in T$ such that $R_{dt} = 1$, as defined for VFSREP(O).

The objective function (O) minimizes the total cost and total weighted amount of extraimmunization of the vaccine formulary subject to the dose requirements for each disease $d \in D$ and extraimmunization restrictions for each disease $d \in D_{NE}$. The objective function weights for extraimmunization (ρ_d and γ_d) are subjective, and hence, allow the model user to weight extraimmunization differently for each disease $d \in D_{NE}$ and/or for time periods when vaccination is permitted versus when vaccination is restricted. However, additional care should be taken in assigning the objective function weights for extraimmunization since the objective function (O) has two objectives (minimize cost and minimize extraimmunization), and hence, given equal emphasis on both of these objectives, the magnitude of the weights for extraimmunization should be scaled similar to the cost parameters c_v, $v \in V$. Constraint (1) ensures that for each disease $d \in D$, at least one vaccine that provides immunization for disease $d \in D$ is administered in some time period when dose $j = 1,2,\ldots,n_d$ may be administered. Constraint (2) and (3) guarantees that for each disease $d \in D$, at least m_{dt} doses of a vaccine that immunize against disease $d \in D$ are administered in the first $t \in T$ time periods, while also ensuring that at most one dose requirement for disease $d \in D$ is satisfied in time period $t \in T$. Constraint (4) and (5) are for each disease $d \in D_{NE}$. Constraint (4) ensures that the total number of doses administered in time periods when vaccination is permitted equals the dose requirement n_d, plus any extra doses that are administered in the time periods when vaccination is permitted. Constraint (5) ensures that the number of doses administered in time periods when vaccination is restricted equals zero, plus any extra doses that are administered in the time periods when vaccination is restricted. Constraint (6), (7), and (8) are the binary and integer constraints on the respective decision variables.

To simplify the formulation of VFSLBREP(O), recall that $T_{dj} = \{t \in T : P_{djt} = 1\}$ is the set of time periods when dose $j = 1,2,\ldots,n_d$, may be administered for disease $d \in D$, where, by assumption, the time periods in T_{dj} are consecutive for all diseases $d \in D$ and doses $j = 1,2,\ldots, n_d$. Furthermore, recall that a disease $d \in D$ is defined to have *mutually exclusive doses* if $T_{di} \cap T_{dj} = \emptyset$ for all $i, j = 1,2,\ldots,n_d$, $i \neq j$ (i.e., the sets T_{dj}, $j = 1,2,\ldots,n_d$ are pairwise mutually exclusive). Note that constraints (2) and (3) are redundant for any disease $d \in D$ with mutually exclusive doses. Furthermore, recall that the variable $Z_{dj}^P \in \mathbf{Z}^+ \cup \{0\}$ is the number of extra vaccine doses administered for disease $d \in D_{NE}$ in all time periods $t \in T$ such that $P_{djt} = 1$, and constraint (4) is also redundant provided that the inequality in constraint (1) becomes an equality for all disease $d \in D_{NE}$ by subtracting the slack variable Z_{dj}^P. Therefore, if every disease has mutually exclusive doses, then VFSLBREP(O) simplifies to the following integer program VFSLBREP(O)-MED (to denote the optimization model where each disease $d \in D$ has mutually exclusive doses).

VFSLBREP(O)-MED

Minimize $\quad \sum_{t \in T} \sum_{v \in V} c_v X_{tv} + \sum_{d \in D_{NE}} \left(\rho_d \left(\sum_{j=1}^{n_d} Z_{dj}^P \right) + \gamma_d Z_d^R \right)$

Subject to

$$\sum_{t \in T} \sum_{v \in V} P_{djt} X_{tv} I_{vd} \geq 1 \qquad \text{for all } d \in D_E,\ j = 1,2,\ldots,n_d,$$

$$\sum_{t \in T} \sum_{v \in V} P_{djt} X_{tv} I_{vd} - Z_{dj}^P = 1 \qquad \text{for all } d \in D_{NE},\ j = 1,2,\ldots,n_d,$$

$$\sum_{t \in T} \sum_{v \in V} R_{dt} X_{tv} I_{vd} - Z_d^R = 0 \qquad \text{for all } d \in D_{NE},$$

$$X_{tv} \in \{0,1\} \qquad \text{for all } t \in T,\ v \in V,$$

$$Z_{dj}^P \in \mathbf{Z}^+ \cup \{0\} \qquad \text{for all } d \in D_{NE},\ j = 1,2,\ldots,n_d,$$

$$Z_d^R \in \mathbf{Z}^+ \cup \{0\} \qquad\qquad \text{for all } d \in D_{NE}.$$

Since all of the diseases in the 2006 Recommended Childhood Immunization Schedule have mutually exclusive doses, the simplification of VFSLBREP(O) to VFSLBREP(O)-MED has practical implications.

Example 14

This extends Example 1 for the childhood immunization schedule depicted in Figure 2. From Example 1, the formulation for VFSLBP(O) (excluding redundant constraints) for this example is:

Minimize $\sum_{t=1}^{8} X_{t1} + 2\sum_{t=1}^{8} X_{t2} + 2\sum_{t=1}^{8} X_{t3} + 3\sum_{t=1}^{8} X_{t4}$

Subject to

$$X_{11} + X_{14} + X_{21} + X_{24} + X_{31} + X_{34} \geq 1$$

$$X_{21} + X_{24} + X_{31} + X_{34} + X_{41} + X_{44} \geq 1$$

$$X_{51} + X_{54} + X_{61} + X_{64} + X_{71} + X_{74} + X_{81} + X_{84} \geq 1$$

$$U_{11} + U_{12} + U_{13} \geq 1$$

$$U_{11} + U_{12} + U_{13} + U_{14} \geq 2$$

$$U_{11} + U_{12} + U_{13} + U_{14} + U_{15} + U_{16} + U_{17} + U_{18} \geq 3$$

$$X_{11} + X_{14} \geq U_{11}$$

$$X_{21} + X_{24} \geq U_{12}$$

$$X_{31} + X_{34} \geq U_{13}$$

$$X_{41} + X_{44} \geq U_{14}$$

$$X_{51} + X_{54} \geq U_{15}$$

$$X_{61} + X_{64} \geq U_{16}$$

$$X_{71} + X_{74} \geq U_{17}$$

$$X_{81} + X_{84} \geq U_{18}$$

$$X_{32} + X_{34} \geq 1$$

$$X_{42} + X_{44} \geq 1$$

$$X_{52} + X_{54} + X_{62} + X_{64} + X_{72} + X_{74} + X_{82} + X_{84} \geq 1$$

$$X_{63} + X_{64} + X_{73} + X_{74} + X_{83} + X_{84} \geq 1$$

$X_{tv} \in \{0,1\}$ for all $t \in T$, $v \in V$

$U_{1t} \in \{0,1\}$ for all $t \in T$.

Suppose that D_{NE} = {2} and $\rho_2 = \gamma_2 = 2$, then for disease d = 2, R_{dt} = 1(0) for time period t = 1,2(3,4,5,6,7,8). Therefore, the formulation for VFSLBREP(O) for this example follows (since disease d = 2 has mutually exclusive doses).

Minimize $\sum_{t=1}^{8} X_{t1} + 2\sum_{t=1}^{8} X_{t2} + 2\sum_{t=1}^{8} X_{t3} + 3\sum_{t=1}^{8} X_{t4} + 2\sum_{j=1}^{3} Z_{2j}^{P} + 2Z_2^R$

Subject to

$$X_{11} + X_{14} + X_{21} + X_{24} + X_{31} + X_{34} \geq 1$$

$$X_{21} + X_{24} + X_{31} + X_{34} + X_{41} + X_{44} \geq 1$$

$$X_{51} + X_{54} + X_{61} + X_{64} + X_{71} + X_{74} + X_{81} + X_{84} \geq 1$$

$$U_{11} + U_{12} + U_{13} \geq 1$$

$$U_{11} + U_{12} + U_{13} + U_{14} \geq 2$$

$$U_{11} + U_{12} + U_{13} + U_{14} + U_{15} + U_{16} + U_{17} + U_{18} \geq 3$$

$$X_{11} + X_{14} \geq U_{11}$$

$$X_{21} + X_{24} \geq U_{12}$$

$$X_{31} + X_{34} \geq U_{13}$$

$$X_{41} + X_{44} \geq U_{14}$$

$$X_{51} + X_{54} \geq U_{15}$$

$$X_{61} + X_{64} \geq U_{16}$$

$$X_{71} + X_{74} \geq U_{17}$$

$$X_{81} + X_{84} \geq U_{18}$$

$$X_{32} + X_{34} - Z_{21}^{P} = 1$$

$$X_{42} + X_{44} - Z_{22}^{P} = 1$$

$$X_{52} + X_{54} + X_{62} + X_{64} + X_{72} + X_{74} + X_{82} + X_{84} - Z_{23}^{P} = 1$$

$$X_{12} + X_{14} + X_{22} + X_{24} - Z_2^R = 0$$

$$X_{63} + X_{64} + X_{73} + X_{74} + X_{83} + X_{84} \geq 1$$

$X_{tv} \in \{0,1\}$ for all $t \in T$, $v \in V$

$U_{1t} \in \{0,1\}$ for all $t \in T$

$Z_{2j}^P, Z_2^R \in \mathbf{Z}^+ \cup \{0\}$ for j = 1,2,3. □

5.2 Computational Complexity and Model Extensions

This section presents the computational complexity of VFSLBREP and describes several model extensions motivated from practical issues in pediatric immunization. Theorem 12 states that VFSLBREP is *NP*-complete.

THEOREM 12: *VFSLBREP is NP-complete in the strong sense.*

PROOF: Let $D_{NE} = \emptyset$, then VFSLBREP becomes VFSLBP, which is *NP*-complete in the strong sense. Furthermore, let $c_v = 0$ for all $v \in V$, then VFSLBREP becomes VFSREP, which is *NP*-complete in the strong sense. ■

The proof of Theorem 12 suggests several special cases of VFSLBREP that remain *NP*-complete. In particular, VFSLBP and VFSREP are both *NP*-complete special cases of VFSLBREP. Therefore, other *NP*-complete special cases of VFSLBREP, as well as special cases that are solvable in polynomial time, are extensions of the cases described in Section 3.2 for VFSLBP and Section 4.2 for VFSREP. For example, Theorem 13 is a natural extension from Section 4.2. Define the linear programming (LP) relaxation of VFSLBREP(O)-MED as the LP model of VFSLBREP(O)-MED along with the relaxed binary variable constraint $0 \le X_{tv} \le 1$ for all time periods $t \in T$ and vaccines $v \in V$ and, for all diseases $d \in D_{NE}$, the relaxed integer variable constraints $Z_{dj}^P \ge 0$ for dose $j = 1,2,\ldots,n_d$, and $Z_d^P \ge 0$.

THEOREM 13: *If Val(v) = 1 for all vaccines $v \in V$, then the LP relaxation of VFSLBREP(O)-MED yields an optimal integer solution.*

PROOF: See the proof of Theorem 9, since the constraint matrix for the LP relaxation of VFSLBREP(O)-MED is equivalent to the constraint matrix for the LP relaxation of VFSREP(O)-MED. ■

There are several formulation extensions of VFSLBREP(O) motivated from practical issues in pediatric immunization. Some of these model extensions apply to any arbitrary childhood immunization schedule while other extensions are specific to a given childhood immunization schedule. Four model extensions are now described. Note that these model extensions also apply to VFSLBP(O) and VFSREP(O) since they are both special cases of VFSLBREP(O).

The first model extension of the formulation for VFSLBREP(O) is for the objective function. Given a flexible childhood immunization schedule (i.e., a childhood immunization schedule where each dose of vaccine may be administered in several time periods (e.g., $|T_{dj}| \geq 2$ for all $d \in D$, $j = 1,2,\ldots,n_d$)), it is likely that there will exist multiple binary variable solutions that are optimal, particularly when the vaccine set contains several monovalent vaccines. For example, consider the constraint $X_{11} + X_{21} + X_{31} \geq 1$ for some disease $d \in D$ and dose 1. Clearly, if vaccine v = 1 is a monovalent vaccine, then the optimal binary variable solution must contain $X_{11} = 1, X_{21} = 0, X_{31} = 0$; $X_{11} = 0, X_{21} = 1, X_{31} = 0$; or $X_{11} = 0$ $X_{21} = 0, X_{31} = 1$. Clearly, the objective function contribution from vaccine v = 1 would be the same regardless if it is administered in time period t = 1, 2, or 3. Likewise, consider the additional constraint $X_{12} + X_{22} + X_{32} \geq 1$ for disease $d' \in D$, dose 1, where vaccine v = 2 is also a monovalent vaccine. In the formulation of VFSLBREP(O) it is optimal to satisfy dose 1 in time period t = 1 for disease $d \in D$ and in time period t = 2 for disease $d' \in D$. Practically speaking, however, it is clearly better to satisfy both doses in time period t = 1, since there are costs associated with each clinic visit. Some of these costs include time for the medical staff, insurance costs and co-pays, and lost wages from time off work for the parent/guardian to attend the clinic visit. Therefore, a practical extension of the formulation for VFSLBREP(O) is to capture the cost of a clinic visit. To model this extension, some additional parameters and decision variables are needed. Define c_C as the cost of a clinic visit and let Y_t = 1 if in time period $t \in T$, a dose of a vaccine is administered to immunize against some disease $d \in D$, 0 otherwise (i.e., Y_t = 1 in time period $t \in T$ if X_{tv} = 1 for some vaccine $v \in V$). Therefore, the objective function for VFSLBREP(O) the captures the cost of a clinic visit is

$$\text{Minimize} \quad \sum_{t\in T} \sum_{v\in V} c_v X_{tv} + \sum_{d\in D_{NE}} \rho_d Z_d^P + \gamma_d Z_d^R + \sum_{t\in T} c_C Y_t$$

subject to constraints (1)-(8) in the formulation of VFSLBREP(O) with the additional constraints

$$X_{tv} \leq Y_t \qquad \text{for all } t \in T, v \in V$$

$$Y_t \in \{0,1\} \qquad \text{for all } t \in T.$$

A second model extension of the formulation for VFSLBREP(O) arises from another practical issue; the ACIP and AAFP currently recommend brand-matching for certain vaccines (CDC 2002). For example, if a child's first dose of vaccine to immunize against diphtheria, tetanus, and pertussis (DTaP) comes from manufacturer m, then each subsequent dose of DTaP vaccine should also come from manufacturer m. To model this extension, some additional notation is needed. Define $D_m \subseteq D$ as the set of diseases requiring brand-matching and partition the vaccine set V by vaccine manufacturers (i.e., $V = \{V_1, V_2, \ldots, V_M\}$, where M is the number of manufacturers). The brand-matching constraints require that a vaccine from the same manufacturer is used for each dose requirement for disease $d \in D_m$. Therefore, for manufacturer $i = 1,2,\ldots,M$, define the following set of constraints

$$\sum_{t \in T} \sum_{v \in V_i} P_{djt} X_{tv} I_{vd} \geq \sum_{t \in T} \sum_{v \in V_i} P_{d(j+1)t} X_{tv} I_{vd} \quad \text{for all } d \in D_m,\ j = 1,\ldots,(n_d - 1).$$

These constraints added to the formulation for VFSLBREP(O) will enforce brand-matching.

The third model extension of the formulation for VFSLBREP(O) is also motivated by recommendations from the ACIP and AAFP. Many vaccines should not be administered until a child has reached a certain age (CDC 2002). For example, a vaccine containing antigens for Hepatitis A should not be administered to children less than one year of age (CDC 2003). To model this extension, an additional parameter is needed. Define

- $t_d \in T$ as the first time period that a vaccine that immunizes against disease $d \in D$ may be administered. Therefore, it is reasonable to assume that for disease $d \in D$, $t_d = \min\{t \in T: P_{djt} = 1 \text{ and } j = 1\}$.

Therefore, the constraints

$$\sum_{t=1}^{t_d} \sum_{v \in V} X_{tv} I_{vd} = 0 \qquad \text{for all } d \in D$$

added to the formulation of VFSLBREP(O) ensure no dose of vaccine is administered for disease $d \in D$ until time period $t_d \in T$.

The final extension of the formulation for VFSLBREP(O) involves separation constraints. Many diseases in the Recommended Childhood Immunization Schedule require a minimum amount of time between each dose requirement of a vaccine based on biological constraints (CDC 2002). These biological separation requirements for disease $d \in D$, dose $j = 1,2,\ldots,n_d$, are enforced by the parameter P_{djt}. For example, in the Recommended

Childhood Immunization Schedule, the first dose of vaccine for *Haemophilus influenzae* type b may be administered at age two months (i.e., $P_{djt} = 1$ for disease d = *Haemophilus influenzae* type b, dose $j = 1$, and time period $t = 3$) and the second dose may be administered at age four months (i.e., $P_{djt} = 1$ for disease d = *Haemophilus influenzae* type b, dose $j = 2$, and time period $t = 4$), which reflects a recommended separation of two months between doses as well as a minimum separation of four weeks (see CDC 2002). However, in recent years, the ACIP and AAFP have simply recommended a vaccine series with recommendations on the number of months between each dose in a series (CDC 2006). For example, the Hepatitis A series was added to the Recommended Childhood Immunization Schedule in 2006, which consists of two doses of vaccine that is recommended for all children at age one year with at least six months between doses (Figure 1 reflects this biological spacing requirement). To model a vaccine series, additional separation constraints may be included in VFSLBREP(O) that capture the biological spacing requirements between doses. To describe these separation constraints, the following parameter is needed in addition to the parameter t_d described above. Let

- $s_d \in \mathbf{Z}^+ \cup \{0\}$ be the number of time periods required between each dose requirement for disease $d \in D$.

Therefore, the separation constraints $\sum_{k=0}^{s_d} \sum_{v \in V} X_{(t+k)v} I_{vd} \leq 1$ for all diseases $d \in D$ and time periods $t = t_d, t_d+1, \ldots, \tau - s_d$ will enforce biological spacing requirements. However, these constraints assume both a *homogeneous* childhood immunization schedule (a childhood immunization schedule where each time period represents the same length of time) and that s_d is independent of the dose $j = 1,2,\ldots,n_d$ for all diseases $d \in D$. Clearly, any non-homogeneous childhood immunization schedule (such as depicted in Figure 1) can be transformed into a homogeneous childhood immunization schedule by adding additional time periods. In practice, however, s_d often changes for each dose requirement. For example, the minimum number of time periods between the third and fourth dose for *Haemophilus influenzae* type b is eight weeks versus the four week minimum separation requirement between the first and second dose (CDC 2002). Therefore, dose j is appended to the parameter s_d, and hence, $s_{dj} \in \mathbf{Z}^+ \cup \{0\}$ is the number of time periods required between dose requirements j and $j+1$ for disease $d \in D$ and $j = 1,2,\ldots,n_d$, where $s_{dn_d} = 0$. This implies the following separation constraints:

$\Sigma_{k=0}^{s_{dj}} \Sigma_{v\in V} X_{(t+k)v} I_{vd} \leq 1$ for all $d \in D$, $j = 1,2,\ldots,n_d$, $t = t_d, t_d+1,\ldots,\tau\text{-}s_{dj}$ such that $P_{djt} = 1$ (SC). These separation constraints result in a more restrictive formulation of VFSLBREP(O).

These separation constraints lead to an interesting complexity result for VFSREP. When separation constraints (SC) are enforced for all diseases $d \in D$, then VFSREP is *NP*-complete for this more restrictive formulation even when $D_{NE} = \emptyset$. This result is stated in Theorem 14.

THEOREM 14: *VFSREP with separation constraints* (*SC*) *is NP-Complete in the strong sense when there is no restriction on extraimmunization* (*i.e.,* $D_{NE} = \emptyset$).

PROOF: First, VFSREP with separation constraints (SC) (VFSREPw/SC) is in *NP* since VFSREP is in *NP* and the separation constraints, $\Sigma_{k=0}^{s_{dj}} \Sigma_{v\in V} X_{(t+k)v} I_{vd} \leq 1$ for all $d \in D$, $j = 1,2,\ldots,n_d$, and $t = t_d, t_d+1,\ldots,\tau\text{-}s_{dj}$ such that $P_{djt} = 1$, may be checked in $O(\upsilon\tau^2\delta)$ time.

To show that VFSREPw/SC is *NP*-complete, a polynomial transformation from 1-in-3 3-SAT with 2-SAT to VFSREPw/SC is constructed.

Given an arbitrary instance of 1-in-3 3-SAT with 2-SAT, define a particular instance of VFSREPw/SC as follows: Set $T = \{1\}$, $D = D_E = \{1,2,\ldots,m+n\}$, $D_{NE} = \emptyset$, $V = \{1,2,\ldots,2n\}$, and $n_1 = n_2 = \cdots = n_{m+n} = 1$. Clearly, $s_{1d} = 0$ and $t_d = 1$ for all $d \in D$. Let the Boolean variables $y_1,y_2,\ldots,y_n$ correspond to vaccines $1,2,\ldots,n$, respectively, and $1\text{-}y_1,1\text{-}y_2,\ldots,1\text{-}y_n$ correspond to vaccines $n+1,n+2,\ldots,2n$, respectively. Let clauses $C_1,C_2,\ldots,C_m$ correspond to diseases $1,2,\ldots,m$, respectively, and $C_{m+1},C_{m+2},\ldots, C_{m+n}$ correspond to diseases $m+1,m+2,\ldots,m+n$, respectively. Set the binary parameters as follows:

$$I_{v=k,d} = \begin{cases} 1 & \text{if the literal } y_k \text{ is in clause } \mathrm{C}_d \text{ for } k = 1,2,\ldots,n; d = 1,2,\ldots,m+n, \text{ respectively} \\ 0 & \text{otherwise,} \end{cases}$$

$$I_{v=(n+k),d} = \begin{cases} 1 & \text{if the literal } (1 - y_k) \text{ is in clause } \mathrm{C}_d \text{ for } k = 1,2,\ldots,n; d = 1,2\ldots,m+n, \text{ respectively} \\ 0 & \text{otherwise.} \end{cases}$$

Therefore, the vaccines that immunize against disease $d = 1,2,\ldots,m+n$, are determined by the literals in clause C_d. Set $P_{djt} = 1$ for all $d \in D$, $j = 1$, $t = 1$, and $Q_{dt} = 1$ for all $d \in D$, $t = 1$. Lastly, $R_{dt} = 0$ for all $d \in D$ since $D_{NE} = \emptyset$. Clearly, this transformation can be made in polynomial time in the size of the arbitrary instance of 1-in-3 3-SAT with 2-SAT. Furthermore, this transformation results in a particular instance of VFSREP where each $d \in D$ has mutually exclusive doses.

To complete the proof, it is necessary to show that a *yes* for this particular instance of VFSREPw/SC implies a *yes* for the arbitrary instance of 1-in-3 3-SAT with 2-SAT, and a *yes* for the arbitrary instance of 1-in-3 3-SAT with 2-SAT implies a *yes* for this particular instance of VFSREPw/SC.

Suppose the answer to the particular instance of VFSREPw/SC is *yes*. The separation constraints require $\Sigma_{v\in V}\ X_{1v}\, I_{vd} \leq 1$ for all $d \in D$ while the dose requirement constraints require $\Sigma_{v\in V}\ X_{1v}\, I_{vd} \geq 1$ for all $d \in D$. Therefore, there must exist values for the binary variables X_{1v}, $v \in V$ such that $\Sigma_{v\in V}\ X_{1v}\, I_{vd} = 1$ for all $d \in D$. Clearly, $I_{vd} = 1$ for $v \in V$, $d \in D$, corresponds to a literal (y_k or 1-y_k for some $k = 1,2,\ldots,n$) that is in clause C_d, $d = 1,2,\ldots,m+n$. Therefore, if $\Sigma_{v\in V}\ X_{1v}\, I_{vd} = 1$ for all $d \in D$, then the binary variable with $X_{1v} = 1$ for each constraint corresponds to the one literal that satisfies clause C_d, for $d = 1,2,\ldots,m+n$. Moreover, $\Sigma_{v\in V}\ X_{1v}\, I_{vd} = 1$ for all $d = m+1, m+2, \ldots, m+n$, and since vaccine k and (n+k) immunize against disease $d = m+k$, the binary variables X_{1k} and $X_{1(n+k)}$ exist together in the constraint for disease $d = m+k$, which implies both y_k and 1-y_k cannot be one for $k = 1,2,\ldots,n$. Therefore, there is a Boolean variable assignment that satisfies all $m+n$ clauses with exactly one true literal, meaning the answer to the arbitrary instance of 1-in-3 3-SAT with 2-SAT is *yes*.

Now suppose the answer to the arbitrary instance of 1-in-3 3-SAT with 2-SAT is *yes*. Then there exists a Boolean variable assignment that results in all $m+n$ clauses being satisfied by exactly one literal. For each Boolean variable y_k, $k = 1,2,\ldots,n$, there are two corresponding binary variables where one such variable (X_{1k}) corresponds to y_k and the other variable ($X_{1(n+k)}$) corresponds to 1- y_k. Therefore, if y_k =1 (0), set X_{1k} = 1 (0) and $X_{1(n+k)}$ = 0 (1). The claim is these values for X_{1v}, $v = 1,2,\ldots,2n$ result in a *yes* answer for the particular instance of VFSREPw/SC. Suppose there does not exist values for the binary variables X_{1v}, $v \in V$, such that $\Sigma_{v\in V}\ X_{1v}\ I_{vd} = 1$ for all $d \in D$. By definition, the constraints for $d = m+1, m+2, \ldots, m+n$ correspond to the k^{th} literal pair y_k and 1-y_k, and hence, if these constraints are not satisfied then neither y_k nor 1-y_k equal one, which is a contradiction. Therefore, for all possible binary variable values of X_{1v}, $v \in V$, there must exist some $d = 1,2,\ldots,m$ such that $\Sigma_{v\in V}\, X_{1v}I_{vd} \neq 1$, which implies $\Sigma_{v\in V}\, X_{1v}I_{vd} = 0$ or $\Sigma_{v\in V}\, X_{1v}I_{vd} > 1$. In either case, the binary variable values of X_{1v} correspond to the Boolean variable values for y_k,

$k = 1,2,\ldots,n$, and I_{vd} identifies the literals in clause C_d for $v = 1,2,\ldots,2n$, $d = 1,2,\ldots,m$. Therefore, if for all possible binary variable values, $\Sigma_{v\in V} X_{1v}I_{vd} = 0$ for some disease $d = 1,2,\ldots,m$, then clause C_d is not satisfied, which contradicts 1-in-3 3-SAT with 2-SAT being *yes*. Likewise, $\Sigma_{v\in V} X_{1v}I_{vd} > 1$ for some disease $d = 1,2,\ldots,m$ implies clause C_d is satisfied by more than one literal, which, again, is a contradiction. Therefore, the values for X_{1v}, $v = 1,2,\ldots,2n$, defined above result in a *yes* answer for the particular instance of VFSREPw/SC.

Furthermore, note that VFSREPw/SC is not a number problem (see Garey and Johnson 1979) since the only numbers occurring in an instance of VFSREPw/SC are the dose requirements n_d for all $d \in D$, which are clearly bounded by $\tau = |T|$, and hence, by the length of the instance. Therefore, VFSREPw/SC is strongly *NP*-complete. ■

5.3 Algorithms and Heuristics

This section discusses how the DP algorithm and the *MAX Rounding* and *Greedy* heuristics for VFSREP(O) presented in Chapter 4 may be extended to VFSLBREP(O). Section 5.3.1 discusses the dynamic programming algorithm for VFSLBREP(O), and Section 5.3.2 discusses the *Rounding*, *MAX Rounding*, and *Greedy* heuristics for VFSLBREP(O)-MED and VFSLBREP(O).

5.3.1 Dynamic Programming Algorithm

This section presents and analyzes a DP algorithm for VFSLBREP(O), which is an extension of the DP presented for VFSREP(O).

Given the stated set of inputs for VFSLBREP(O), the DP algorithm solves VFSLBREP(O) one period at a time beginning at the first time period (i.e., $t = 1$), and steps through each time period in T until $t = \tau$. The parameters, sets, states, state space, decisions and decision space for the DP algorithm for VFSLBREP(O) are those previously defined for VFSREP(O) in Section 4.3.1.

Given that decision $\mathbf{Y}_t = \mathbf{S}_t - \mathbf{S}_{t-1}$, then a transition from state $\mathbf{S}_{t-1} \in \Omega_{t-1}$ to state $\mathbf{S}_t \in \Omega_t$ requires that a dose of vaccine be administered in time period $t \in T$ for each disease in the set $D_t = \{d \in D : Y_{td} = 1\}$. The sets $V_t = \{v \in V: I_{vd} = 1$ and $d \in D_t\}$ (i.e., the set of vaccines that immunize against any disease that requires vaccination in time period $t \in T$) and D_t define a sub-instance of VFSLBREP(O), termed SCP-IP($\mathbf{Y}_t$). To describe SCP-IP($\mathbf{Y}_t$), recall the following definitions:

- $D_{Et} = D_E \cap D_t$ and $D_{NEt} = D_{NE} \cap D_t$ for any time period $t \in T$,
- $Z^P_{dt} \in \mathbf{Z}^+ \cup \{0\}$ be the number of extra doses of vaccine administered for disease $d \in D_{NE}$ in time period $t \in T$ such that $Y_{td} = 1$,
- $Z^R_{dt} \in \mathbf{Z}^+ \cup \{0\}$ be the number of extra doses of vaccine administered for disease $d \in D_{NE}$ in time period $t \in T$ such that $Y_{td} = 0$, (i.e., for disease $d \in D_{NE} \backslash D_{NEt}$).

The specific sub-instance for VFSLBREP(O) for time period $t \in T$ and decision $\mathbf{Y}_t \in \Phi_t$ is given by

SCP-IP($\mathbf{Y}_t$)

$$\text{Minimize} \quad \sum_{v \in V_t} c_v X_{tv} + \sum_{d \in D_{NEt}} \rho_d Z^P_{dt} + \sum_{d \in D_{NE} \backslash D_{NEt}} \gamma_d Z^R_{dt}$$

Subject to

$$\sum_{v \in V_t} X_{tv} I_{vd} \geq 1 \qquad \text{for all } d \in D_{Et},$$

$$\sum_{v \in V_t} X_{tv} I_{vd} - Z^P_{dt} = 1 \qquad \text{for all } d \in D_{NEt},$$

$$\sum_{v \in V_t} X_{tv} I_{vd} - Z^R_{dt} = 0 \qquad \text{for all } d \in D_{NE} \backslash D_{NEt},$$

$$X_{tv} \in \{0,1\} \qquad \text{for all } v \in V_t,$$

$$Z^P_{dt}, Z^R_{dt} \in \mathbf{Z}^+ \cup \{0\} \qquad \text{for all } d \in D_{NE}.$$

To characterize the cost of decision $\mathbf{Y}_t \in \Phi_t$, which is the cost of transitioning from state $\mathbf{S}_{t-1} \in \Omega_{t-1}$ in time period $(t-1) \in T$ to state $\mathbf{S}_t \in \Omega_t$ in time period $t \in T$, define the one-period cost function $C_t(\mathbf{S}_{t-1}, \mathbf{Y}_t)$ as the value of the objective function in time period $t \in T$ given state $\mathbf{S}_{t-1} \in \Omega_{t-1}$ and decision $\mathbf{Y}_t \in \Phi_t$. Note, however, that this one-period cost in time period $t \in T$ depends only on decision $\mathbf{Y}_t \in \Phi_t$, and hence, the optimal value of SCP-IP($\mathbf{Y}_t$) = $C_t(\mathbf{S}_{t-1}, \mathbf{Y}_t) = C_t(\mathbf{Y}_t)$. Therefore, the optimal one-period value over all possible decisions in time period $t \in T$ is given by $\min_{\mathbf{Y}_t \in \Phi_t} C_t(\mathbf{Y}_t)$.

Define $Z_t(\mathbf{S}_t)$ as the minimum value of the objective function of a vaccine formulary that immunizes against all diseases through time period $t \in T$ subject to the number of required doses at the end of time period $t \in T$ being equal to $\mathbf{S}_t \in \Omega_t$. Therefore, the DP optimality equation is given by the recurrence relation

$$Z_t(\mathbf{S}_t) = \min_{\mathbf{Y}_t \in \Phi_t, \mathbf{S}_{t-1} \in \Omega_{t-1}: \mathbf{S}_t = \mathbf{S}_{t-1} + \mathbf{Y}_t} \{C_t(\mathbf{Y}_t) + Z_{t-1}(\mathbf{S}_{t-1})\}.$$

Furthermore, the minimum value of the objective function that satisfies a given childhood immunization schedule is given by

$$z^* = \min_{\mathbf{S}_\tau \in \Omega_\tau} Z_t(\mathbf{S}_t),$$

where Ω_τ is the state space for the final time period $\tau \in T$. The DP algorithm for VFSLBREP(O) is now formally given.

Dynamic Programming Algorithm for VFSLBREP(O)

Step 1. Initialize:

a. Initial state, $\mathbf{S}_0 \leftarrow \mathbf{0}$ (the δ-dimensional zero vector)
b. Initial objective function contribution, $Z_0(\mathbf{S}_0) \leftarrow 0$
c. Set m_{d0}, $M_{d0} \leftarrow 0$ for all $d \in D$
d. Initial stage, $t \leftarrow 1$

Step 2. Compute

$$Z_t(\mathbf{S}_t) = \min_{\mathbf{Y}_t \in \Phi_t, \mathbf{S}_{t-1} \in \Omega_{t-1}: \mathbf{S}_t = \mathbf{S}_{t-1} + \mathbf{Y}_t} \{C_t(\mathbf{Y}_t) + Z_{t-1}(\mathbf{S}_{t-1})\}$$

for each state $\mathbf{S}_t \in \Omega_t$.

Step 3. If $t < \tau$, then $t \leftarrow t + 1$ and return to Step 2. Else, stop and return $z^* = \min_{\mathbf{S}_\tau \in \Omega_\tau} Z_t(\mathbf{S}_t)$.

To determine the complexity of this DP algorithm, suppose that the SCP-IP($\mathbf{Y}_t$) problem instance with δ diseases and υ vaccines can be solved in $O(\mathbf{T}_{\text{SCP-IP}})$ time. Furthermore, define $\mathbf{S}_{Max}$ to be the maximum number of possible states within any time period $t \in T$. Each time period requires $O((\mathbf{S}_{Max})^2 \cdot \mathbf{T}_{\text{SCP-IP}})$ time, and hence, with τ time periods, the DP algorithm for VFSLBREP(O) executes in $O(\tau(\mathbf{S}_{Max})^2 \cdot \mathbf{T}_{\text{SCP-IP}})$ time. The DP algorithm's worst case complexity may be improved, however, since each SCP-IP($\mathbf{Y}_t$) instance depends only on the decision vector $\mathbf{Y}_t \in \Phi_t$. Therefore, SCP-IP($\mathbf{Y}_t$) for decision $\mathbf{Y}_t \in \Phi_t$ only needs to be solved once. It can be shown that the complexity of solving for all possible decisions is $O(\upsilon\delta 2^\delta)$. This means that for each time period $t \in T$, the complexity of Step 2 becomes $O(\delta(\mathbf{S}_{Max})^2)$, and hence, the DP algorithm has a $O(\tau\delta(\mathbf{S}_{Max})^2 + \upsilon\delta 2^\delta)$ worst case time complexity, which is an improvement over $O(\tau(\mathbf{S}_{Max})^2 \cdot \mathbf{T}_{\text{SCP-IP}})$ when $\mathbf{S}_{Max}$ is large. To exploit this added efficiency, the implementation of the DP algorithm may employ a 'branch and remember' recursive algorithm to find the optimal value for each SCP-IP($\mathbf{Y}_t$) instance. Therefore, SCP-IP($\mathbf{Y}_t$) need only be computed once using the recursive algorithm *SCP-IP*.

This recursive algorithm assumes $\rho_d = \gamma_d$ for all diseases $d \in D_{NE}$. Initially, the given set of diseases for $\mathbf{Y}_t$ is D_t, and hence, $D' = D_t$.

SCP-IP(D')
If $D' = \emptyset$, return 0 as the solution value
If SCP-IP($\mathbf{Y}_t$) for $D' = \{d \in D : Y_{td} = 1\}$ has been solved previously, return its optimal value
Select a disease $d \in D'$ that requires immunization
Let $V' = \{v \in V: I_{vd} = 1\}$ (set of vaccines $v \in V$ that immunize against disease $d \in D'$)
Set $BestValue = +\infty$
For each vaccine $v \in V'$
Let $D^* = D' \setminus \{d \in D': I_{vd} = 1\}$
$Value = SCP\text{-}IP(D^*)$ (find the optimal penalty for the set of diseases D^*)
Let $D_{NEv} = \{d \in \overline{D'} : d \in D_{NE}, I_{vd} = 1\}$
Set $Penalty_v = 0$
For each disease $d \in D_{NEv}$
$Penalty_v = Penalty_v + \gamma_d$
If $Value + c_v + Penalty_v < BestValue$
$BestValue = Value + c_v + Penalty_v$
Store $BestValue$ for D' (save the optimal solution for the set of diseases D')
Return $BestValue$

Despite its exponential worst case complexity run time, the DP algorithm for VFSLBREP(O) offers several advantages as described in Section 3.3.1 for VFSLBP(O).

5.3.2 Heuristics

This section discusses how the *Rounding*, *MAX Rounding,* and *Greedy* heuristics for VFSREP(O) presented in Chapter 4 may be extended to VFSLBREP(O).

Both the *Rounding* and *MAX Rounding* heuristics use the solution from a linear program (LP) to construct a feasible binary solution. Relaxing the binary and integer constraints for the decision variables in VFSLBREP(O)-MED yields the LP relaxation

$$\text{Minimize} \quad \sum_{t\in T}\sum_{v\in V} c_v X_{tv} + \sum_{d\in D_{NE}} \left(\rho_d \left(\sum_{j=1}^{n_d} Z_{dj}^P\right) + \gamma_d Z_d^R\right)$$

Subject to

$$\sum_{t\in T}\sum_{v\in V} P_{djt} X_{tv} I_{vd} \geq 1 \qquad \text{for all } d \in D_E,\ j = 1,2,\ldots,n_d,$$

$$\sum_{t\in T}\sum_{v\in V} P_{djt} X_{tv} I_{vd} - Z_{dj}^P = 1 \qquad \text{for all } d \in D_{NE},\ j = 1,2,\ldots,n_d,$$

$$\sum_{t\in T}\sum_{v\in V} R_{dt} X_{tv} I_{vd} - Z_d^R = 0 \qquad \text{for all } d \in D_{NE},$$

$$0 \leq X_{tv} \leq 1 \qquad \text{for all } t \in T,\ v \in V,$$

$Z_{dj}^{P} \geq 0$ for all $d \in D_{NE}$, $j = 1,2,\ldots,n_d$,

$Z_{d}^{R} \geq 0$ for all $d \in D_{NE}$.

Recall that X_{LP}^{*} denotes the optimal decision vector for the LP relaxation of VFSLBREP(O)-MED and $X_{LP\,tv}^{*}$, $t \in T$, $v \in V$, Z_{dj}^{P*}, $d \in D_{NE}$, j=1,2,…,n_d, and Z_{d}^{R*}, $d \in D_{NE}$, denote the optimal values for the decision variables in the LP relaxation. Likewise, recall that $\alpha_d \equiv (\Sigma_{v\in V} I_{vd})(\max_{j=1,2,\ldots,n_d} \Sigma_{t\in T} P_{djt})$ for all diseases $d \in D$ and $\alpha \equiv \max_{d\in D} \alpha_d$. The *Rounding* heuristic for VFSLBREP(O)-MED is then equivalent to the *Rounding* heuristic for VFSREP(O)-MED and is now is now formally given.

Rounding Heuristic for VFSLBREP(O)-MED

Step 1. Solve the LP relaxation of VFSLBREP(O)-MED
Step 2. $X_{tv} \leftarrow 0$ for all $t \in T$ and $v \in V$
Step 3. For all $t \in T$ and $v \in V$
- a. If $X_{LP_{tv}}^{*} \geq 1/\alpha$, then $X_{tv} \leftarrow 1$

Step 4. For all $d \in D_{NE}$
- a. For all $j = 1,2,\ldots,n_d$
  - i. $Z_{dj}^{P} \leftarrow \Sigma_{t\in T}\Sigma_{v\in V} P_{djt} X_{tv} I_{vd} - 1$
- b. $Z_{d}^{R} \leftarrow \Sigma_{t\in T}\Sigma_{v\in V} R_{dt} X_{tv} I_{vd}$

Step 5. Compute and return $\sum_{t\in T}\sum_{v\in V} c_v X_{tv} + \sum_{d\in D_{NE}}\left(\rho_d\left(\sum_{j=1}^{n_d} Z_{dj}^{P}\right) + \gamma_d Z_{d}^{R}\right)$.

Lemma 10 establishes the feasibility of the solution returned by the *Rounding* heuristic.

LEMMA 10: *The Rounding heuristic for VFSLBREP(O)-MED returns a feasible binary solution X,* (*i.e., a decision vector that satisfies the childhood immunization schedule*).

PROOF: Follows directly from the proof of Lemma 9. ■

Given that linear programming is solvable in polynomial time, it then follows that the *Rounding* heuristic executes in polynomial time. Theorem 15 shows that the value of the binary solution returned by the *Rounding* heuristic is guaranteed to be no worse than $\alpha \cdot z_{IP}$.

THEOREM 15: *The Rounding heuristic is an α-approximation algorithm for VFSREP*(*O*)*–MED.*

PROOF: Clearly, the *Rounding* heuristic executes in polynomial time since LP executes in polynomial time. Recall that the optimal objective function values of VFSLBREP(O)-MED and its LP relaxation are denoted by z_{IP} and z_{LP}, respectively, where $z_{LP} \leq z_{IP}$ (since the feasible region of VFSLBREP(O)-MED is contained in the feasible region of its LP relaxation). It remains to show that $\sum_{t\in T}\sum_{v\in V} c_v X_{tv} + \sum_{d\in D_{NE}} (\rho_d \sum_{j=1}^{n_d} Z_{dj}^P + \gamma_d Z_d^R) \le \alpha \cdot z_{IP}$. By step 4 of the algorithm,

$$\sum_{t\in T}\sum_{v\in V} c_v X_{tv} + \sum_{d\in D_{NE}} (\rho_d \sum_{j=1}^{n_d} Z_{dj}^P + \gamma_d Z_d^R) = \sum_{t\in T}\sum_{v\in V} c_v X_{tv} + \sum_{d\in D_{NE}} (\rho_d \sum_{j=1}^{n_d}(\sum_{t\in T}\sum_{v\in V} P_{djt} X_{tv} I_{vd} - 1) + \gamma_d \sum_{t\in T}\sum_{v\in V} R_{dt} X_{tv} I_{vd})$$

$$\le \sum_{t\in T}\sum_{v\in V} c_v\, (X_{LP_{tv}}^* \alpha) + \sum_{d\in D_{NE}} \sum_{t\in T}\sum_{v\in V}(\rho_d \sum_{j=1}^{n_d} (P_{djt}\, (X_{LP_{tv}}^* \alpha) I_{vd} - 1) + \gamma_d R_{dt}\, (X_{LP_{tv}}^* \alpha) I_{vd})$$

(since $X_{tv} = 1$ only if $X_{LP_{tv}}^* \alpha \ge 1$)

$$= \alpha \sum_{t\in T}\sum_{v\in V} c_v\, X_{LP_{tv}}^* + \alpha \sum_{d\in D_{NE}} \sum_{t\in T}\sum_{v\in V}(\rho_d \sum_{j=1}^{n_d}(P_{djt}\, X_{LP_{tv}}^* I_{vd} - 1) + \gamma_d R_{dt}\, X_{LP_{tv}}^* I_{vd})$$

$$= \alpha(\sum_{t\in T}\sum_{v\in V} c_v\, X_{LP_{tv}}^* + \sum_{d\in D_{NE}} (\rho_d \sum_{j=1}^{n_d}(\sum_{t\in T}\sum_{v\in V} P_{djt}\, X_{LP_{tv}}^* I_{vd} - 1) + \gamma_d \Sigma_{t\in T}\Sigma_{v\in V} R_{dt} X_{LP_{tv}}^* I_{vd}))$$

$$= \alpha(\sum_{t\in T}\sum_{v\in V} c_v\, X_{LP_{tv}}^* \sum_{d\in D_{NE}} (\rho_d \sum_{j=1}^{n_d} Z_{dj}^{P*} + \gamma_d Z_d^{R*}))$$

$$= \alpha \cdot z_{LP}$$

$$\le \alpha \cdot z_{IP} \quad (\text{since } z_{LP} \le z_{IP}). \ \blacksquare$$

Theorem 15 implies that Corollaries 7 and 8 for VFSREP(O)-MED presented in Section 4.3.2 are also valid for VFSLBREP(O)-MED.

If X_{LP}^* contains several fractional variables, then the *Rounding* heuristic tends to round too many variables to one, thereby yielding a significant amount of extraimmunization. Instead of rounding all variables greater than or equal to the $1/\alpha$ threshold, it seems reasonable to round only a few variables with large fractional values (i.e., variables closest to one), since these variables are more likely to equal one in the optimal solution. The *MAX Rounding* heuristic limits the number of rounded variables by selecting the variables with large fractional values.

To present the *MAX Rounding* heuristic, recall the notation $\mathbf{D} = \{(d, j): d \in D, j = 1,2,\ldots,n_d\}$ to be the set of all diseases ordered by dose, where $|\mathbf{D}| = \Sigma_{d=1}^{\delta} n_d$, and, for all time periods $t \in T$ and vaccines $v \in V$, $C_{tv} = \{(d, j) \in \mathbf{D}: I_{vd} = 1 \text{ and } P_{djt} = 1\}$, which specifies the diseases and dose that vaccine $v \in V$ immunizes against in time period $t \in T$. Therefore, $C_{tv} \subseteq \mathbf{D}$ for all time periods $t \in T$ and vaccines $v \in V$. Furthermore, in the case when all diseases $d \in D$ have mutually exclusive doses, at most one $(d, j) \in \mathbf{D}$ for all diseases $d \in D$ is contained in any set C_{tv}, since for a given disease $d \in D$ and time period $t \in T$, $P_{djt} = 1$ for at most one dose $j = 1,2,\ldots,n_d$, and hence, each set C_{tv} does not contain multiple doses for any disease $d \in D$. Lastly, recall that $f_{tv} = X^*_{LP_{tv}}$ for all time periods $t \in T$ and vaccines $v \in V$, which specifies the value of vaccine $v \in V$ in time period $t \in T$. Therefore, the *MAX Rounding* heuristic limits the number of rounded variables by greedily selecting (at each iteration) the most valuable available vaccine $v \in V$ that immunizes against the most disease doses (not yet covered) in time period $t \in T$ (i.e., rounds the variable $X^*_{LP_{tv}}$ that maximizes $f_{tv} \cdot |C_{tv}|$) until every disease dose $(d, j) \in \mathbf{D}$ is covered by some vaccine $v \in V$ in time period $t \in T$. The *MAX Rounding* heuristic is now formally given.

MAX Rounding Heuristic for VFSLBREP(O)-MED

Step 1. Initialize:

a. Solve the LP relaxation of VFSLBREP(O)-MED
b. $f_{tv} \leftarrow X^*_{LP_{tv}}$ for all $t \in T$, $v \in V$ such that $X^*_{LP_{tv}} \geq 1/\alpha$
c. $X_{tv} \leftarrow 0$ for all $t \in T$ and $v \in V$
d. $\hat{C}_{tv} \leftarrow C_{tv}$ for all $t \in T$ and $v \in V$

Step 2. While $\mathbf{C} = \bigcup_{\{tv: X_{tv}=1\}} C_{tv} \neq \mathbf{D}$ do

a. $(t', v') \leftarrow \arg\max_{t \in T, v \in V} f_{tv} \cdot |\hat{C}_{tv}|$ (select the non-empty set $\hat{C}_{tv}$ with the largest fractional value times the number of disease doses covered by vaccine $v \in V$ in time period $t \in T$)
b. $X_{t'v'} \leftarrow 1$ (administer vaccine $v' \in V$ in time period $t' \in T$)
c. $\hat{C}_{tv} \leftarrow \hat{C}_{tv} \setminus \hat{C}_{t'v'}$ for all $t \in T$ and $v \in V$ (remove all the disease doses covered by vaccine $v' \in V$ in time period $t' \in T$ from all remaining sets)

Step 3. For all $d \in D_{NE}$

a. For all $j = 1,2,\ldots,n_d$
   i. $Z^P_{dj} \leftarrow \Sigma_{t \in T}\Sigma_{v \in V} P_{djt} X_{tv} I_{vd} - 1$
b. $Z^R_d \leftarrow \Sigma_{t \in T}\Sigma_{v \in V} R_{dt} X_{tv} I_{vd}$

Step 4. Compute and return $\sum_{t \in T} \sum_{v \in V} c_v X_{tv} + \sum_{d \in D_{NE}} \left(\rho_d \left(\sum_{j=1}^{n_d} Z^P_{dj}\right) + \gamma_d Z^R_d\right)$.

The *MAX Rounding* heuristic executes in $O(\mathbf{T}_{\text{LP}} + |\mathbf{D}|\tau\upsilon)$ time, where $\mathbf{T}_{\text{LP}}$ is the time required to solve the LP relaxation of VFSLBREP(O)-MED. Furthermore, the *MAX Rounding* heuristic returns a feasible solution, since every iteration of the while loop (i.e., Step 2) administers a vaccine that satisfies at least one dose requirement for some disease $d \in D$ (i.e., every iteration covers at least one $(d, j) \in \mathbf{D}$). Moreover, Step 1.b. ensures that the solution returned by the *MAX Rounding* heuristic can be no worse than the solution returned by the *Rounding* heuristic, and hence, the *MAX Rounding* heuristic is also an α-approximation algorithm for VFSLBREP(O)-MED.

Like the *Rounding* and *MAX Rounding* heuristics, the *Greedy* heuristic for VFSREP(O)-MED is easily adapted to VFSLBREP(O)-MED. As before, the *Greedy* heuristic iteratively selects the vaccine with the lowest cost that also incurs the smallest penalty for extraimmunization and immunizes against the most disease doses. Recall that the extraimmunization penalty for vaccine $v \in V$ in time period $t \in T$ is $W_{tv} = \sum_{\{d \in D: I_{vd}=1\}} w_{dt}$, where

$$w_{dt} = \begin{cases} \rho_d & \text{if } d \in D_{NE}, (d, j) \in C_{tv} \text{ for some } j = 1,2,\ldots,n_d \text{, and } (d, j) \in \mathbf{C} = \bigcup_{\{tv: X_{tv}=1\}} C_{tv} \\ \gamma_d & \text{if } d \in D_{NE}, (d, j) \notin C_{tv} \text{ for some } j = 1,2,\ldots,n_d \\ 0 & \text{otherwise,} \end{cases}$$

Therefore, for VFSLBREP(O)-MED, the total "cost" of vaccine $v \in V$ in time period $t \in T$ is $c_v + W_{tv}$, and the *Greedy* heuristic is now formally given.

Greedy Heuristic for VFSLBREP(O)-MED

Step 1. Initialize:
  a. $X_{tv} \leftarrow 0$ for all $t \in T$ and $v \in V$
  b. $\hat{C}_{tv} \leftarrow C_{tv}$ for all $t \in T$ and $v \in V$

Step 2. While $\mathbf{C} = \bigcup_{\{tv: X_{tv}=1\}} C_{tv} \neq \mathbf{D}$ do
  a. Compute W_{tv} for all $t \in T$ and $v \in V$ (compute extraimmunization penalty for vaccine $v \in V$ in time period $t \in T$)
  b. $(t', v') \leftarrow \arg\min_{t \in T, v \in V} (c_v + W_{tv}) / |\hat{C}_{tv}|$ (select the non-empty set $\hat{C}_{tv}$ with the smallest "cost" per disease doses covered by vaccine $v \in V$ in time period $t \in T$. Break ties by selecting vaccine $v \in V$ that immunizes against the most diseases in time period $t \in T$.)
  c. $X_{t'v'} \leftarrow 1$ (administer vaccine $v' \in V$ in time period $t' \in T$)
  d. $\hat{C}_{tv} \leftarrow \hat{C}_{tv} \setminus \hat{C}_{t'v'}$ for all $t \in T$ and $v \in V$ (remove all the disease doses covered by vaccine $v' \in V$ in time period $t' \in T$ from all remaining sets)

Step 3. For all $d \in D_{NE}$

a. For all $j = 1,2,\ldots,n_d$

i. $Z^P_{dj} \leftarrow \Sigma_{t\in T}\Sigma_{v\in V} P_{djt} X_{tv} I_{vd} - 1$

b. $Z^R_d \leftarrow \Sigma_{t\in T}\Sigma_{v\in V} R_{dt} X_{tv} I_{vd}$

Step 4. Compute and return $\sum_{t\in T}\sum_{v\in V} c_v X_{tv} + \sum_{d\in D_{NE}}\left(\rho_d\left(\sum_{j=1}^{n_d} Z^P_{dj}\right) + \gamma_d Z^R_d\right)$.

The *Greedy* heuristic executes in $O(|\mathbf{D}|\tau\upsilon)$ time, and returns a feasible solution, since every iteration of the while loop (i.e., Step 2) administers a vaccine that satisfies at least one dose requirement for some disease $d \in D$ (i.e., every iteration covers at least one $(d, j) \in \mathbf{D}$). Therefore, the *Greedy* heuristic should (in practice) be more efficient than the *MAX Rounding* heuristic.

Finally, a *MAX Rounding* and *Greedy* heuristics for VFSLBREP(O) is performed by converting a VFSLBREP(O) instance into two distinct VFSLBREP(O)-MED instances, and then applying the *MAX Rounding* and *Greedy* heuristics for each VFSLBREP(O)-MED instance to find a feasible solution for the VFSLBREP(O) instance as described in Section 4.3.5 for VFSREP(O).

Chapter 6: Conclusion and Extensions

This chapter concludes the dissertation and presents several possible research extensions that are of practical and theoretical interest.

6.1 Conclusion

This dissertation applied operations research methodologies to designing pediatric vaccine formularies that capture the combinatorial explosion of alternatives and choices spawn by combination vaccines and ensure that immunity is safely achieved by restricting or limiting extraimmunization. In particular, the dissertation presented three fundamental problems for designing pediatric vaccine formularies.

The first problem, VFSLBP, extended the research discussed in Chapter 2 by formulating a general decision problem and discrete optimization problem that seeks a minimum cost vaccine formulary for a given childhood immunization schedule. The second model, VFSREP, examined the issue of extraimmunization in pediatric immunization by formulating a general decision problem and discrete optimization problem that seek a vaccine formulary for a given childhood immunization schedule that restricts extraimmunization. The final problem, VFSLBREP, combines VFSLBP and VFSREP by formulating a general decision problem and discrete optimization problem that seeks a minimum cost vaccine formulary for a given childhood immunization schedule that also restricts extraimmunization. As more combination vaccines come to market and the Recommended Childhood Immunization Schedule becomes more complex to include more diseases and cover more time periods, these problems will capture the combinatorial explosion of alternatives for public health policy-makers and administrators, vaccine manufacturers, pediatricians, and parents/guardians by identifying minimum cost vaccine formularies and schedules that safely use combination vaccines, which will help address safety concerns in pediatric immunization, reduce costs, and reduce vaccine wastage associated with extraimmunization.

In general, VFSLBP(O), VFSREP(O), and VFSLBREP(O) were shown to be *NP*-hard unless the vaccines, schedule parameters, or disease set are significantly restricted. Therefore, the existence of an algorithm that finds the optimal solution in polynomial time is unlikely, unless $P = NP$. This dissertation presented a DP algorithm that solves VFSLBP(O), VFSREP(O), and VFSLBREP(O) to optimality, but becomes intractable as the size of the childhood immunization schedule grows, particularly as the size of the disease set grows.

In Section 3.3.1, this DP algorithm for VFSLBP(O) was compared computationally to an IP B&B algorithm. These results showed that the DP algorithm was significantly more efficient (at least eight times faster) when the size of the disease set is reasonable. However, for most of the randomly generated childhood immunization schedules both the DP and IP B&B algorithms required at least an order of magnitude more time to execute when compared to the execution time of the heuristics presented in Section 3.3. The *MAX Rounding* and *Greedy* heuristics returned a cost within ten percent of the optimal solution (on average) for each set of test problems. Moreover, the average execution time for each heuristic was less sensitive to increases in the size of the childhood immunization schedule or the valency of the vaccine set.

In Section 4.3.1, this DP algorithm for VFSREP(O) was compared computationally to an IP B&B algorithm, namely IP-MIN. These results showed that the DP algorithm was significantly more efficient (approximately 300 to 1000 times faster) when the size of the disease set is reasonable. Furthermore, the execution time of the DP algorithm was insensitive to the size of the set D_{NE} (the average execution time remained nearly constant at 1.74 seconds when δ_{NE} = 4 to 1.81 seconds when δ_{NE} = 11), whereas the execution time of IP-MIN was sensitive to the size of the set D_{NE} (the average execution time tripled from 518 seconds when δ_{NE} = 4 to 1767 seconds when δ_{NE} = 11). However, for most of the randomly generated childhood immunization schedules, both the DP and IP B&B algorithms required at least twice as much time to execute when compared to the execution time of the heuristics presented in Section 4.3. Moreover, the average execution time for each heuristic was less sensitive to increases in the size of the childhood immunization schedule or the valency of the vaccine set.

The heuristics presented in this dissertation will allow more efficient analysis of larger childhood immunization schedules and practical analysis involving Monte Carlo simulation or finding an optimal vaccine formulary for each child on a case-by-case basis, which will require the solution of several unique VFSLBP(O), VFSREP(O), or VFSLBREP(O) instances.

6.2 Research Extensions

There are several avenues to further extend the research reported in this dissertation. First, further research is needed to determine how the model extensions (motivated by real-world constraints) described in Section 5.2 affect the computational complexity and the performance and feasibility of the DP algorithms and heuristics for VFSLBP(O), VFSREP(O), and VFSLBREP(O). Furthermore, alternative objective functions or cost parameters might also be considered, such as minimizing the number of injections in each time period, minimizing the number of time periods required to satisfy the childhood immunization schedule, or considering a quadractic objective function. Next, the robust structure of the DP algorithm is well suited for adding uncertainty (like that described for the balking problem in Section 3.3.1) to VFSLBP(O), VFSREP(O), and VFSLBREP(O), which would make this research even more practical to the pediatric public health community.

Other research of theoretical interest and importance is to determine or strengthen approximation bounds provided for the heuristics presented here, and to explore non-approximability for VFSLBP(O), VFSREP(O), and VFSLBREP(O). For example, the approximation bounds for the heuristics for VFSLBP(O)-MED do not apply to the solution returned by the *A* heuristic for VFSLBP(O), since the *A* heuristic for VFSLBP(O) converts a VFSLBP(O) instance into two distinct VFSLBP(O)-MED instances. Moreover, developing new heuristics with better approximation bounds, added efficiency, and/or improved empirical results is also of interest.

Finally, extending the problems described in this dissertation to other applications areas in health care or other industries may provide practical and useful insights. For example, consider the Air Force Acquisition with Limited Budget Problem (AFALBP). To describe this problem, some additional background is needed. The United States Department of Defense (DoD) has a resource allocation process called the Planning, Programming, and Budgeting System (PPBS). As a department within DoD, the United States Air Force participates in this process. With billions of dollars to allocate across personnel, weapon systems, base infrastructures, and military operations, the budgeting process is quite daunting. Within the Department of the Air Force, there exist several major commands (MAJCOMS) that are responsible to train, organize, and equip the force. Each MAJCOM participates in the PPBS process by first developing a long-term strategic master plan, then using this plan, each MAJCOM programs a six year resource allocation proposal subject to fiscal constraints. Over the past decade, several MAJCOMS have begun using optimization models to make better (more quantitative and justifiable) resource allocation decisions. Typically, each MAJCOM has a mission that is defined by several key tasks. The MAJCOM develops and procures weapon systems such as satellites, fighters, bombers, and ballistic missiles to accomplish these key mission tasks. Therefore, an important objective for each MAJCOM is to cover these mission tasks by designing a strategic plan for weapon system development/procurement within a limited budget environment. This problem, termed the Air Force Acquisition with Limited Budget Problem, asks whether it is possible to design a procurement schedule that covers the key mission tasks within a fiscally constrained budget and shares the same fundamental structure as VFSLBP. This problem is now formally stated.

Air Force Acquisition with Limited Budget Problem (AFALBP)

Given:

- A set of time periods, $T = \{t_1, t_2, \ldots, t_{|T|}\}$,
- a set of mission tasks, $M = \{m_1, m_2, \ldots, m_{|M|}\}$,
- a set of systems $S = \{s_1, s_2, \ldots, s_{|S|}\}$ available for procurement to provide coverage for the $|M|$ mission tasks,
- the number of coverage periods that a system must be procured for coverage of $|M|$ mission tasks, $n_1, n_2, \ldots, n_{|M|}$,
- the cost of the $|S|$ systems, $c_1, c_2, \ldots, c_{|S|}$,
- a budget B,
- a set of binary parameters that indicate which systems cover which mission tasks. Therefore, $I_{smt} = 1$ if system s provides coverage for mission task m in time period t, 0 otherwise, for $s \in S$, $m \in M$, $t \in T$.
- a set of binary parameters that indicate the set of time periods that a system may be procured to cover a mission task for a particular coverage period; therefore, $P_{mjt} = 1$ if in time period $t \in T$, a system that covers mission task m may be procured to provide coverage for the j^{th} coverage period, 0 otherwise, for $m \in M$, $j = 1,2,\ldots,n_m$, $t \in T$.
- a set of binary parameters that indicate the set of time periods in which a system may be administered to satisfy any coverage requirement for mission task m; therefore, $Q_{mt} = 1$ if in time period $t \in T$, a system may be procured to satisfy any coverage requirement

against mission task $m \in M$, 0 otherwise, (i.e., for any mission task $m \in M$ and time period $t \in T$, $Q_{mt} = 1$ if and only if $P_{mjt} = 1$ for some coverage period $j = 1,2,\ldots,n_m$.),

- a set of integer parameters that indicate the minimum number of coverage periods a system is required for mission task $m \in M$ through time period $t \in T$; denoted m_{mt}.

Question: Does there exist a set of systems from S that can be procured over time periods T such that these systems provide coverage against all the mission tasks in M, at a total cost no greater than B (i.e., do there exist values for the binary variables X_{ts}, $t \in T$, $s \in S$, where $X_{ts} = 1$ if system s is procured in time period t, 0 otherwise, such that for all $m \in M$, $\Sigma_{t \in T}\Sigma_{s \in S}\, P_{mjt}\, X_{ts}\, I_{smt} \geq 1$ for $j = 1,2,\ldots,n_m$ and $\Sigma_{t=1,2,\ldots,t'}\Sigma_{s \in S}\, Q_{mt}\, X_{ts}\, I_{smt} \geq m_{mt'}$ for all time periods $t' \in T$ and $\Sigma_{t \in T}\Sigma_{s \in S}\, c_s\, X_{ts} \leq B$)?

Clearly, the computational complexity, algorithms, and heuristics for VFSLBP may be extended to AFALBP.

Bibliography

Ahuja, R.K., Magnanti, T.L., and Orlin, J.B. 1993. *Network Flows*. Prentice-Hall: Upper Saddle River, New Jersey.

Asano, T. and Williamson, D.P. 2002. Improved Approximation Algorithms for MAX SAT. *Journal of Algorithms*; 42:173-202.

Bazaraa, M.S., Jarvis, J.J., and Sherali, H.D. 1990. *Linear Programming and Network Flows*. John Wiley & Sons: New York.

Binder, S., Levitt, A.M., Sacks, J.J., and Hughes. 1999. Emerging Infectious Diseases: Public Health Issues for the 21st Century. *Science*; 284(5418):1311-1313.

Burden, R.L. and Faires, J.D. 1997. *Numerical Analysis*, 6th ed. Brookes-Cole: New York.

Caprara, A., Toth, P., and Fischetti, M. 2000. Algorithms for the Set Covering Problem. *Annals of Operations Research*; 98:353-371.

CDC Vaccine Price List. 2005. Center for Disease Control and Prevention, National Immunization Program Website. November 21, 2005, http://www.cdc.gov/nip/vfc/cdc_vac_price_list.htm.

Centers for Disease Control and Prevention. 2006. Recommended Childhood and Adolescent Immunization Schedule—United States. *Morbidity and Mortality Weekly Report*; 54(52):Q1-Q4.

Centers for Disease Control and Prevention. 2005. Recommended Childhood and Adolescent Immunization Schedule—United States. *Morbidity and Mortality Weekly Report*; 53:Q1-Q3.

Centers for Disease Control and Prevention. 2003. Prevention and Control of Infections with Hepatitis Viruses in Correctional Settings. Appendix Hepatitis A and B Vaccines. *Morbidity and Mortality Weekly Report*; 52(RR-1);34-36.

Centers for Disease Control and Prevention. 2002. General Recommendations on Immunization. *Morbidity and Mortality Weekly Report*; 51(RR-2).

Centers for Disease Control and Prevention. 1999. Combination Vaccines for Childhood Immunization. *Morbidity and Mortality Weekly Report*; 48(18)(RR-5).

Centers for Disease Control and Prevention. 1995. Recommended Childhood and Adolescent Immunization Schedule—United States. *Morbidity and Mortality Weekly Report*; 44(RR-5);1-9.

Chen R.T., DeStefano, F., Pless R., Mootrey, G., Kramarz, P. and Hibbs, B. 2001. Challenges and Controversies in Immunization Safety. *Infectious Disease Clinics of North America*; 15(1):21-39.

Cochi, S.L. 2005. *National Immunization Survey*. Presentation to National Press Club, July 27, 2005.

Cohen, M.L. 2000. Changing patterns in infectious disease. *Nature*; 406:762-767.

Diekema, D.S. 2005. Responding to Parental Refusals of Immunization of Children. *Pediatrics*; 115(5):1428-1431.

Edwards, K.M. and Decker, M.D. 2001. Combination Vaccines. *Infectious Disease Clinics of North America*: *Vaccine Recommendations, Challenges and Controversies*; 15(1):209-230.

Flanagan-Klygis, E.A, Sharp, L., and Frader, J.E. 2005. Dismissing the Family Who Refuses Vaccines: A Study of Pediatrician Attitudes. *Archives of Pediatrics & Adolescent Medicine*; 159(10):929-934.

Garey, M.R. and Johnson, D.S. 1979. *Computers and Intractability*: *A Guide to the Theory of NP-Completeness*. W.H. Freeman: New York.

Hochbaum, D.S., ed. 1997. *Approximation Algorithms for NP-hard Problems*. PWS Publishing Company. Boston, Massachusetts.

Hoffman, K. and Padberg, M. 2006. Set Covering, Packing, and Partitioning Problems. Last retrieved March 27, 2006, from http://iris.gmu.edu/~khoffman/papers/set_covering.html

Infectious Diseases in Children. 2002. Almost 200 new drugs in development for use in children. 15(10):41-44.

Jacobson, S.H., Karnani, T. and Sewell, E.C. 2003a. Analyzing the economic value of the hepatitis B-*Haemophilus influenzae* type B combination vaccine by reverse engineering a formulary selection algorithm. *Vaccine*; 21:2169-2177.

Jacobson, S.H., Karnani, T. and Sewell, E.C. 2004. Assessing the Impact of Wastage on Pediatric Vaccine Immunization Formulary Costs Using a Vaccine Selection Algorithm. *Vaccine*; 22:2307-2315.

Jacobson, S.H. and Sewell, E.C. 2002. Using Monte Carlo Simulation to Determine Combination Vaccine Price Distributions for Childhood Disease. *Health Care Management Science*; 5:135-145.

Jacobson, S.H., Sewell, E.C., Allwine, D.A., Medina, E.A. and Weniger, B.G. 2003b. Designing pediatric vaccine formularies and pricing pediatric combination vaccines using operations research models and algorithms. *Expert Rev. Vaccines*; 2(1):15-19.

Jacobson, S.H., Sewell, E.C., Deuson, R., and Weniger, B.G. 1999. An integer programming model for vaccine procurement and delivery for childhood immunization: a pilot study. *Health Care Management Science*; 2:1-9.

Jacobson, S.H., Sewell, E.C., and Karnani, T. 2005. Engineering the Economic Value of Two Pediatric Combination Vaccines. *Health Care Management Science*; 8(1):29-40.

Jacobson, S.H., Sewell, E.C., Proano, R.A., and Jokela, J.A. 2006a. Stockpile levels for pediatric vaccines: How much is enough? *Vaccine*; 24:3530-3537.

Jacobson, S.H., Sewell, E.C., Proano, R.A. 2006b. An Analysis of the Pediatric Vaccine Shortage Problem. *Health Care Management Science*; (to appear).

Mackay, I.R. and Rosen, F.S. 2001. Advances in Immunology: Vaccines and Vaccination. *New England Journal of Medicine*; 345(14):1042-1053.

McGuire, T.G. 2003. Setting Prices for New Vaccines (in Advance). *International Journal of Health Care Finance and Economics*; 3:207-224.

Miltersen, P.B. 2005. Algorithms for Satisfiability. Course Notes for Combinatorial Search, April 27, 2005. Version 4.1. Retrieved from http://www.daimi.au.dk/dKS/sat2005.pdf.

National Vital Statistics Reports. 2004. 53(6):33-34.

Nemhauser, G.L. and Wolsey, L.A. 1999. *Integer and Combinatorial Optimization*. John Wiley & Sons, New York.

Plotkin, S.A. and Orenstein, W.A. 2004. *Vaccines*, 4th ed. Elsevier Inc, USA.

Rotwani, R. and Raghavan P. 1995. *Randomized Algorithms*. Cambridge University Press, United Kingdom.

Sewell, E.C. and Jacobson, S.H. 2003. Using an Integer Programming Model to Determine the Price of Combination Vaccines for Childhood Immunization. *Annals of Operations Research*; 119:261-284.

Sewell, E.C., Jacobson, S.H. and Weniger, B.G. 2001. Reverse engineering a formulary selection algorithm to determine the economic value of pentavalent and hexavalent combination vaccines. *Pediatric Infectious Disease Journal*; 20(11):S45-S56.

Sewell, E.C., Stuart K., Jacobson, S.H., and Hall, S.N. 2005. Selecting Optimal Pediatric Vaccine Formularies Using a Dynamic Programming Algorithm. Technical Report, Southern Illinois University, Edwardsville, IL.

Tawfik, S. 2005. Mixed-Integer LP. MATLAB Central File Exchange Website. Retrieved Oct 4, 2005, from http://www.mathworks.com/matlabcentral/fileexchange/loadCategory. .do.

Weniger, B.G. 1996. Economic analysis to meet the challenges of new combination vaccines, Presented at *Vaccine Economics*: *Planning a Research Agenda for the Challenge of New and Improved Vaccines*, Atlanta, GA.

Weniger, B.G., Chen, R.T., Jacobson, S.H., Sewell, E.C., Deuson, R., Livengood, J.R., and Orenstein, W.A. 1998. Addressing the challenges to immunization practice with an economic algorithm for vaccine selection. *Vaccine*; 16(19):1885-1887.

Author's Biography

Captain Shane N. Hall, United States Air Force, graduated from Show Low High School in Show Low, Arizona in May 1991, and then attended Brigham Young University in Provo, Utah. He received a Bachelor of Science degree in mathematics and his officer commission in April 1997. After receiving his commission, Shane entered active military service as a scientific analyst and worked in the 96th Communications Group at Eglin AFB, Florida. In August 1998, he entered the Graduate School of Engineering and Management, Air Force Institute of Technology, Wright-Patterson AFB, Ohio and received his Master of Science degree in operations research in March 2000. After receiving his Masters, he worked as an analyst in the Office of Aerospace Studies at Kirtland AFB, New Mexico until August 2003 when he began his doctoral studies in the Department of Mechanical and Industrial Engineering.

While studying the design and analysis of pediatric vaccine formularies under the guidance of Professor Sheldon H. Jacobson, Shane has co-authored two book chapters, six journal publications, and presented his research at professional conferences. He also received the MIE George W. Harper award for excellence in safety engineering.